



Average frequencies of digits in infinite IFS's and applications to continued fractions and Lüroth expansions

L. Olsen¹ · M. West²

Received: 17 December 2018 / Accepted: 3 August 2020
© The Author(s) 2020

Abstract

The detailed investigation of the distribution of frequencies of digits of points belonging to attractors K of Infinite iterated functions systems (IIFS's) is a fundamental and important problem in the study of attractors of IIFS's. This paper studies the Baire category of different families of sets of points belonging to attractors of IIFS's characterised by the behaviour of the frequencies of their digits. All our results are of the following form:

a typical (in the sense of Baire) point $x \in K$ has the following property: the average frequencies of digits of x have maximal oscillation.

We consider general types of average frequencies, namely, average frequencies associated with general averaging systems. These averages include, for example, all higher order Hölder and Cesaro averages, and Riesz averages. Surprising, for *all* averaging systems (regardless of how powerful they are) we prove that a typical (in the sense of Baire) point $x \in K$ has the following property: the average frequencies of digits of x have maximal oscillation. This substantially extends previous results and provides a powerful topological manifestation of the fact that “points of divergence” are highly visible. Several applications are given, e.g. to continued fraction digits and Lüroth expansion digits.

Keywords Baire category · Non-normal numbers · Average systems · Infinite iterated function systems · Continued fraction expansion · Lüroth expansion · Hausdorff dimension · Packing dimension

Mathematics Subject Classification Primary 28A80

Communicated by Adrian Constantin.

✉ L. Olsen
lo@st-and.ac.uk

Extended author information available on the last page of the article

1 Introduction

Self-similar sets are arguably the most important examples of fractal sets. This is due to their simple description and the fact that they exhibit many of the properties one expects from fractals. Self-similar sets also form a natural and fundamental framework for studying frequencies of digits associated with a variety of classical and important digit expansions from number theory. Indeed, a unique feature of self-similar sets is that points belonging to a self-similar set can be assigned a sequence of digits (the precise definitions will be given in Sect. 2). For example, for judiciously chosen self-similar sets the assigned digits of a real number x in the unit interval $[0, 1]$ coincide with the usual N -ary digits of x or the continued fraction expansion digits of x , and the assigned digits therefore form a natural generalisation of the classical notions of digits associated with traditional number theoretical expansions.

The detailed investigation of the distribution of digits is a fundamental and important problem in the study of self-similar sets. For example, the detailed knowledge of the distribution of digits is crucial for our understanding of dimensions of self-similar fractals. Indeed, lower bounds for the Hausdorff and packing dimensions of self-similar sets are invariably obtained by using the ergodic theorem to examine the limiting distribution of almost all frequencies of digits (more precisely, to find the local dimension of judiciously constructed ergodic measures); these results are then translated into lower bounds for the fractal dimension using the “distribution of mass” principle. The detailed study of the distribution of digits also links to fundamental questions in metric number theory. In this direction, the detailed study of almost sure frequencies of digits of different expansions is a fundamental area in metric number theory with deep connections to, for example, Diophantine approximations, see [14] and [5]. Divergence points, i.e. points for which the frequencies of digits, or more general ergodic averages, do not converge, have also been studied. Indeed, until very recently divergence points have been considered of little interest in dynamical systems and geometric measure theory and according to folklore, sets of divergence points carried no essential information about the underlying structure. However, recently Barrière & Schmeling [3] have shown that many natural and important sets of divergence points associated with self-similar constructions are highly “visible” or “observable”, namely, they have full Hausdorff dimension. Sets of divergence points have been investigated further in, for example, [26] showing that these sets have a surprising rich and intricate fractal structure. A common feature in most studies of convergence and divergence properties of frequencies of digits (and, in particular, in the work cited above), is that they focus exclusively on the frequencies of digits of *almost all* points chosen with respect to judiciously constructed ergodic measures.

However, there is a further fundamental notion of “generic”, namely, the notion of *typical* defined using Baire category. Recall that if X is a metric space and P is a property that the elements of X may have, then we say that a typical element x in X has property P if the set $E = \{x \in X \mid x \text{ does not have property } P\}$ is of the 1st Baire category (also sometimes called meagre), see Oxtoby [27] for more details. While properties of distributions of digits of *almost all* points have been investigated intensively, the study of distributions of digits of *typical* points is sporadic and focusing on specific and concrete examples. For example, some basic questions of the distribution

of the usual N -ary digits of typical numbers in the unit interval have been investigated (see, for example, [1,12,25,31] and the references therein), and some basic questions of the distribution of digits associated with other types of expansions, including the continued fraction expansion and the Lüroth expansion, of typical numbers in the unit interval have been investigated (see, for example, [2,17–19,29,30] and the references therein). However, somewhat surprisingly, a general and comprehensive study of the distributions of digits of *typical* points seems to be missing, and the main purpose of this paper is to provide a detailed and comprehensive investigation of this problem in general and abstract setting. More precisely, we focus on points belonging to self-similar sets generated by general *Infinite Iterated Functions Systems* using the powerful (and classical) notion of *Averaging Systems*. We will now explain this in more detail.

Infinite Iterated Functions Systems Loosely speaking, self-similar sets come in two varieties, namely, those associated with a (finite) Iterated Functions System (IFS), and those associated with an Infinite Iterated Functions System (IIFS). The theory of IFS's go back to Moran's paper from 1946 [22] and found its present elegant and mature formulation in Hutchinson's seminal paper [11] from 1981. The more general notion of IIFS is more recent and was introduced and pioneered by Mauldin & Urbanski in late 1990's [20,21]. For a positive integer N , the N -ary digits of a real number coincide with the digits associated with a particular (finite) IFS. However, the digits of other important expansions, including, for example, the continued fraction expansion and the Lüroth expansion, do not coincide with the digits associated with any (finite) IFS. Instead, the continued fraction digits and the Lüroth expansion digits coincide with the digits associated with two judiciously constructed IIFS's. For this reason, we investigate digits of points belonging to arbitrary IIFS's satisfying the so-called Strong Open Set Condition (and 2 further very mild technical assumptions). We emphasize that we do not assume that the IIFS we study consist of conformal maps. In fact, the maps may be highly non-conformal. We also stress that this setup includes a large number of classical expansions, including, for example, the continued fraction expansion and the Lüroth expansion. Indeed, as an application of our main results we provide a detailed study of the frequencies of the continued fraction expansion digits and the Lüroth expansion digits of typical real numbers; this is done in Sects. 5.1 and 5.2, respectively. Short studies of the distribution digits of typical points belonging to self-similar sets associated with IFS's and IIFS's are present in [2,17,18]. However, while the results in [2,17,18] are formulated in the setting of IFS's or IIFS's, the main focus in [2,17,18] is on N -ary expansion of real numbers and on the continued fraction expansions of irrational numbers.

Averaging systems Using a Baire category viewpoint, often frequencies do not exist and, in fact, diverges in the most dramatic way possible. To explore this further we analyse very general types of averages of frequencies using the classical notion of an Averaging System. This notion is very general and includes many classical averages, including, for example, *all* higher order Hölder averages and *all* higher order Cesaro averages.

Surprising, for *all* averaging systems (regardless of how powerful they are) we prove that a typical (in the sense of Baire) point x has the following property: the average frequencies of digits of x have maximal oscillation. For example, it follows

from this that a typical (in the sense of Baire) point x has the following property: *all* higher order Hölder averages and *all* higher order Cesaro averages of digits of x have maximal oscillation. This substantially extends previous results and provides a powerful topological manifestation of the fact that “points of divergence” are highly visible.

The structure of the paper is the following. We state the main results in Sect. 2. As an application of our main results we provide a detailed study of higher order Hölder averages and higher order Cesaro averages of frequencies of digits of typical points in Sect. 3, and as a further application, we study the logarithmic Riesz averages of frequencies of digits of typical points in Sect. 4. Finally, specialising further, frequencies of the continued fraction expansion digits and the Lüroth expansion digits of typical real numbers are investigated in Sects. 5.1 and 5.1, respectively. The proofs are presented in Sects. 5–8.

2 Statement of results

2.1 Infinite iterated function systems

In this section we define the notion of an Infinite Iterated Function System and the associated self-similar set.

Definition *Infinite Iterated Function System* An Infinite Iterated Function System (IIFS) in \mathbb{R}^d is a list $(X, (S_i)_{i \in I})$ where I is a set with $1 < |I| \leq \aleph_0$ (here $|I|$ denotes the cardinality of I) and

- (1) X is a compact subset of \mathbb{R}^d with $\overline{\bigcirc} X = X$;
- (2) For each $i \in I$, the map $S_i : X \rightarrow X$ is an injective contraction and there is a constant $0 < s < 1$ such that $|S_i(x) - S_i(y)| \leq s|x - y|$ for all $i \in I$ and all $x, y \in X$.

If the set I is finite, then the list $(X, (S_i)_{i \in I})$ is called an Iterated Function System (IFS).

Next, we define the self-similar set (also sometimes called the limit set) associated with an IIFS $(X, (S_i)_{i \in I})$. We first introduce the following notation. For a positive integer n , let

$$I^n = \{\mathbf{i} = i_1 \dots i_n \mid i_j \in I \text{ for all } j\}.$$

i.e. I^n is the family of all strings $\mathbf{i} = i_1 \dots i_n$ of length n with entries $i_j \in I$. Also, let

$$I^* = \bigcup_n I^n,$$

$$I^{\mathbb{N}} = \{\mathbf{i} = i_1 i_2 \dots \mid i_j \in I \text{ for all } j\},$$

i.e. I^* is the family of all finite strings $\mathbf{i} = i_1 \dots i_n$ with entries $i_j \in I$ and $I^{\mathbb{N}}$ is the family of all infinite strings $\mathbf{i} = i_1 i_2 \dots$ with entries $i_j \in I$; here and below \mathbb{N} denotes

the set of all strictly positive integers, i.e. $\mathbb{N} = \{1, 2, \dots\}$. For $\mathbf{i} = i_1 i_2 \dots \in I^{\mathbb{N}}$ and $n \in \mathbb{N}$, write $\mathbf{i}|n = i_1 \dots i_n$ for the truncation of \mathbf{i} at the n 'th place. Finally, for $\mathbf{i} = i_1 \dots i_n \in I^*$, we write

$$S_{\mathbf{i}} = S_{i_1} \cdots S_{i_n}$$

Since clearly $\text{diam}(S_{\mathbf{i}|n}(X)) \leq s^n \text{diam}(X)$ for $\mathbf{i} \in I^{\mathbb{N}}$, it follows that $(S_{\mathbf{i}|n}(X))_n$ is decreasing sequence of non-empty compact sets, and the intersection $\bigcap_n S_{\mathbf{i}|n}(X)$ is therefore a singleton. In particular, we can define the map $\pi : I^{\mathbb{N}} \rightarrow X$ by

$$\{\pi(\mathbf{i})\} = \bigcap_n S_{\mathbf{i}|n}(X). \quad (2.1)$$

The self-similar set K associated with the IIFS $(X, (S_i)_{i \in I})$ is now defined by

$$K = \pi(I^{\mathbb{N}}), \quad (2.2)$$

see, for example, [20,21]. It is not difficult to see that the limit set K satisfies the following “self-similar” identity

$$K = \bigcup_{i \in I} S_i(K). \quad (2.3)$$

If the set I is finite, then it is well-known (see, for example [11] or the textbooks [7,8]) that K is the unique non-empty and compact subset of \mathbb{R}^d that satisfies the “self-similar” identity (2.3); however, if I is infinite, then the set K may not be compact. We will often assume that the IIFS $(X, (S_i)_{i \in I})$ satisfies the following separation condition; loosely speaking, this condition says that overlaps $S_i(K) \cap S_j(K)$ are small for $i \neq j$.

Definition *The Strong Open Set Condition* An IIFS $(X, (S_i)_{i \in I})$ is said to satisfy the Strong Open Set Condition (SOSC) if there is a non-empty open set $V \subseteq X$ with $V \cap K \neq \emptyset$ such that $S_i(V) \cap S_j(V) = \emptyset$ for all $i, j \in I$ with $i \neq j$.

The reader is referred to the Falconer's excellent textbook [8] for a discussion of the dimension theory of self-similar sets associated with IFS's and to Mauldin & Urbanski's text [21] for a detailed discussion of the dimension theory of self-similar sets associated with IIFS's.

2.2 Frequencies of digits

If $x \in K$ and $x = \pi(\mathbf{i})$ where $\mathbf{i} = i_1 i_2 \dots \in I^{\mathbb{N}}$, then we refer to the i_j 's as digits of x ; we note that a point $x \in K$ may have several strings of digits and a string of digits for x is therefore not, in general, unique. To investigate the divergence and convergence properties of the frequencies of digits of x we introduce the following

notation. Let $\mathbf{i} = i_1 i_2 \dots \in I^{\mathbb{N}}$ be an infinite string. For a positive integer n and a finite string $\mathbf{l} = l_1 \dots l_k \in I^*$, we write

$$\pi_n(\mathbf{i}; \mathbf{l}) = \frac{|\{1 \leq j \leq n \mid i_j = l_1, \dots, i_{j+k-1} = l_k\}|}{n} \quad (2.4)$$

for the frequency of the string \mathbf{l} among the first n digits of the string \mathbf{i} . We also let

$$\boldsymbol{\pi}(\mathbf{i}; \mathbf{l}) = (\pi_n(\mathbf{i}; \mathbf{l}))_n \quad (2.5)$$

denote the sequence of frequencies $\pi_n(\mathbf{i}; \mathbf{l})$. Finally, we define the lower and upper frequencies of the string \mathbf{l} among the digits of the string \mathbf{i} by

$$\underline{\pi}(\mathbf{i}; \mathbf{l}) = \liminf_n \pi_n(\mathbf{i}; \mathbf{l})$$

and

$$\overline{\pi}(\mathbf{i}; \mathbf{l}) = \limsup_n \pi_n(\mathbf{i}; \mathbf{l}).$$

In addition, write

$$\boldsymbol{\pi}_{k,n}(\mathbf{i}) = (\pi_n(\mathbf{i}; \mathbf{l}))_{\mathbf{l} \in I^k}, \quad (2.6)$$

i.e. $\boldsymbol{\pi}_{k,n}(\mathbf{i})$ denotes the vector of frequencies $\pi_n(\mathbf{i}; \mathbf{l})$ of all strings \mathbf{l} of length equal to k .

2.3 Typical frequencies of digits

We define the subset Δ_k of $\ell^1(I^k)$ by

$$\Delta_k = \left\{ (p_{\mathbf{l}})_{\mathbf{l} \in I^k} \mid p_{\mathbf{l}} \geq 0, \sum_{\mathbf{l}} p_{\mathbf{l}} = 1 \right\},$$

i.e. Δ_k denotes the simplex of probability vectors indexed by strings $\mathbf{i} = i_1 \dots i_k \in I^k$ of length k . We will always equip Δ_k with the 1-norm $\|\cdot\|_1$. The vector $\boldsymbol{\pi}_{k,n}(\mathbf{i})$ clearly belongs to Δ_k . We will now quantify the divergence of the frequencies of the digits of the point $x = \pi(\mathbf{i})$ by considering the extent to which the sequence $(\boldsymbol{\pi}_{k,n}(\mathbf{i}))_n$ fills up the simplex Δ_k . Of course, in general, it is not true that the sequence $(\boldsymbol{\pi}_{k,n}(\mathbf{i}))_n$ fills up a substantial part of Δ_k for any \mathbf{i} . For example, let $I = \mathbb{N}$ and consider strings of length 3. By considering all possible ways a string of length 2, such as 37 $\in \mathbb{N}^2$ (i.e. 37 represents the string of length 2 whose first digit equals 3 and whose second digit equals 7), can arise it is easily seen that

$$\left| \sum_{l \in \mathbb{N}} \pi_n(\mathbf{i}; l37) - \sum_{l \in \mathbb{N}} \pi_n(\mathbf{i}; 37l) \right| \leq \frac{1}{n}$$

for all $\mathbf{i} \in \mathbb{N}^{\mathbb{N}}$. This implies that for each $\mathbf{i} \in \mathbb{N}^{\mathbb{N}}$, all but finitely many points in the sequence $(\pi_{3,n}(\mathbf{i}))_n$ will be very close to the subsimplex

$$\Delta_3 \cap \left\{ (x_{\mathbf{l}})_{\mathbf{l} \in \mathbb{N}^3} \mid \sum_{\mathbf{l} \in \mathbb{N}} x_{\mathbf{l}37} = \sum_{\mathbf{l} \in \mathbb{N}} x_{37\mathbf{l}} \right\}. \quad (2.7)$$

Hence, in general the sequence $(\pi_{k,n}(\mathbf{i}))_n$ will not fill up a significant part of the simplex Δ_k , and the full simplex Δ_k is not the “correct” object to consider. Rather we need to consider the subsimplex defined by slicing Δ_k by various planes corresponding to the subsimplex in (2.7). Motivated by this, we define the subsimplex Δ_k^{shift} of shift invariant probability vectors in \mathbb{R}^{I^k} by

$$\Delta_k^{\text{shift}} = \left\{ (p_{\mathbf{l}})_{\mathbf{l} \in I^k} \mid p_{\mathbf{l}} \geq 0, \sum_{\mathbf{l}} p_{\mathbf{l}} = 1, \sum_{\mathbf{l}} p_{\mathbf{l}\mathbf{i}} = \sum_{\mathbf{l}} p_{\mathbf{l}\mathbf{u}} \text{ for all } \mathbf{i} \in I^{k-1} \right\}. \quad (2.8)$$

The next result shows that the subsimplex Δ_k^{shift} of shift invariant probability vectors is, indeed, the correct subsimplex to consider. To make this statement precise, we introduce the following notation, namely, if $(x_n)_n$ is a sequence in a metric space X , we will write $\text{acc}_n x_n$ for the set of accumulation points of $(x_n)_n$, i.e.

$$\text{acc}_n x_n = \left\{ x \in X \mid x \text{ is an accumulation point of } (x_n)_n \right\}.$$

We can now state the result that says that the subsimplex Δ_k^{shift} is the correct subsimplex to consider; in Theorem A and in all subsequent parts of the paper we will always equip Δ_k^{shift} with the 1-norm and all accumulation points are with respect to the 1-norm.

Theorem A [18,23] *Let $\mathbf{i} \in I^{\mathbb{N}}$. Then $\text{acc}_n \pi_{k,n}(\mathbf{i}) \subseteq \Delta_k^{\text{shift}}$.*

In view of Theorem A it is natural to quantify the divergence of the sequence $\pi_{k,n}(\mathbf{i}) = (\pi_n(\mathbf{i}; \mathbf{I}))_{\mathbf{l} \in I^k}$ by considering the extent to which it fills up the subsimplex Δ_k^{shift} . We therefore introduce the following notation. For $\mathbf{i} \in I^{\mathbb{N}}$, we say that the point $\pi(\mathbf{i})$ is extremely non- k -normal if the set of accumulation points of the sequence equals Δ_k^{shift} , and we will denote the set of extremely non- k -normal points by \mathbb{E}_k , i.e.

$$\mathbb{E}_k = \pi \left\{ \mathbf{i} \in I^{\mathbb{N}} \mid \text{acc}_n \pi_{k,n}(\mathbf{i}) = \Delta_k^{\text{shift}} \right\}.$$

We will say that a point is extremely non-normal if it is extremely non- k -normal for all k and we let \mathbb{E} denote the set of extremely non-normal, i.e.

$$\mathbb{E} = \bigcap_k \mathbb{E}_k.$$

Hence, the points in \mathbb{E} are those whose frequencies of digits diverge in the worst possible way. Work by [2,17,18,23], generalising earlier results [12,25,29–31], shows

somewhat surprisingly that, in some cases, the set \mathbb{E} of extremely non-normal points is extremely big from a topological viewpoint; these results are summarised in Theorem B below.

Theorem B [2,17,18,23] *Let $(X, (S_i)_{i \in I})$ be an IIFS and assume that one of the following two conditions (a) or (b) is satisfied:*

- (a) *The set I is finite and the list $(X, (S_i)_{i \in I})$ satisfies the SOSC;*
- (b) *The list $(X, (S_i)_{i \in I})$ is the continued fraction IIFS, i.e. $I = \mathbb{N}$, $X = [0, 1]$ and $S_i : [0, 1] \rightarrow [0, 1]$ is defined by $S_i(x) = \frac{1}{x+i}$ for all i .*

Then the set \mathbb{E} is comeager in K .

2.4 Typical average frequencies of digits

If a sequence diverges, forming the sequence of averages may “smoothen out” the irregularities of the original sequence. We will now use this idea to study of the irregular behaviour of the sequence of frequencies $\pi(\mathbf{i}; \mathbf{I}) = (\pi_n(\mathbf{i}; \mathbf{I}))_n$. Indeed, Theorem B shows that, in some cases, the sequence of frequencies $\pi(\mathbf{i}; \mathbf{I}) = (\pi_n(\mathbf{i}; \mathbf{I}))_n$ diverges (in the worst possible way) for a typical point $\pi(\mathbf{i})$ in K , and it is therefore natural to form the sequence of average frequencies and investigate the convergence of this sequence. In fact, below we form very general types of averages, namely, averages formed using the classical notion of an averaging (or summability) system, and show (perhaps surprisingly) that *all* of these averages diverge in the worst possible way. We start by recalling the definition of an averaging (or summability) system; the reader is referred to Hardy’s classical text [9] for a systematic treatment of averaging systems.

Definition *Averaging system* An averaging system is a sequence $\Pi = (\Pi_n)_{n \in \mathbb{N}}$ of sequences $\Pi_n = (\pi_{n,i})_{i \in \mathbb{N}}$ of real numbers satisfying the following conditions:

- (1) $\pi_{n,i} \geq 0$ for all n and i ;
- (2) the set $\{i \mid \pi_{n,i} > 0\}$ is finite for all n ;
- (3) $\frac{1}{n} \sum_i \pi_{n,i} \rightarrow 0$;
- (4) The Consistency Condition: if $(x_n)_n$ is a sequence of positive real numbers such that there is a real number x with $x_n \rightarrow x$, then $\sum_i \pi_{n,i} x_i \rightarrow x$.

For a sequence $\mathbf{x} = (x_n)_n$ of positive real numbers and a positive integer n , write

$$A_{\Pi,n}(\mathbf{x}) = \sum_i \pi_{n,i} x_i,$$

and define the lower and upper Π -average of \mathbf{x} by

$$\underline{A}_{\Pi}(\mathbf{x}) = \liminf_n A_{\Pi,n}(\mathbf{x})$$

and

$$\overline{A}_{\Pi}(\mathbf{x}) = \limsup_n A_{\Pi,n}(\mathbf{x}),$$

respectively.

Remark We immediately note that the usual Cesaro average of a sequence can be obtained by considering the averaging system defined by $\Pi = (\Pi_n)_n$ where $\Pi_n = (\frac{1}{n}, \dots, \frac{1}{n}, 0, 0, \dots)$ (where the sequence Π_n contains n terms of the form $\frac{1}{n}$); indeed, with for this choice of Π , we clearly have $A_{\Pi,n}(\mathbf{x}) = \frac{1}{n}(x_1 + \dots + x_n)$ for any sequence $\mathbf{x} = (x_n)_n$. More general averages, including, for example, all higher order Cesaro and Hölder averages and the logarithmic Riesz averages can also be obtained from suitable averaging systems and will be studied in detail in Sects. 3 and 4, respectively.

Applying averaging systems to the sequence $\boldsymbol{\pi}(\mathbf{i}; \mathbf{l}) = (\pi_n(\mathbf{i}; \mathbf{l}))_n$ of frequencies leads to the definition of average frequencies.

Definition *Average frequencies* Let $\Pi = (\Pi_n)_n$ be an averaging system with $\Pi_n = (\pi_{n,i})_{i \in \mathbb{N}}$. Let $\mathbf{i} \in I^{\mathbb{N}}$. For a finite string $\mathbf{l} \in I^k$, we define the lower and upper Π -average frequency of the string \mathbf{l} among the digits of the string \mathbf{i} by

$$\underline{\pi}_{\Pi}(\mathbf{i}; \mathbf{l}) = \underline{A}_{\Pi}(\boldsymbol{\pi}(\mathbf{i}; \mathbf{l}))$$

and

$$\overline{\pi}_{\Pi}(\mathbf{i}; \mathbf{l}) = \overline{A}_{\Pi}(\boldsymbol{\pi}(\mathbf{i}; \mathbf{l})).$$

I.e. if we write

$$\pi_{\Pi,n}(\mathbf{i}; \mathbf{l}) = A_{\Pi,n}(\boldsymbol{\pi}(\mathbf{i}; \mathbf{l})) = \sum_i \pi_{n,i} \pi_i(\mathbf{i}; \mathbf{l}),$$

then

$$\underline{\pi}_{\Pi}(\mathbf{i}; \mathbf{l}) = \liminf_n \pi_{\Pi,n}(\mathbf{i}; \mathbf{l}) = \liminf_n \sum_i \pi_{n,i} \pi_i(\mathbf{i}; \mathbf{l})$$

and

$$\overline{\pi}_{\Pi}(\mathbf{i}; \mathbf{l}) = \limsup_n \pi_{\Pi,n}(\mathbf{i}; \mathbf{l}) = \limsup_n \sum_i \pi_{n,i} \pi_i(\mathbf{i}; \mathbf{l}).$$

Remark We note that the usual frequencies $\pi_n(\mathbf{i}; \mathbf{l})$ are average frequencies; indeed, if we define the averaging system by $\Pi = (\Pi_n)_n$ where $\Pi_n = (\delta_{n,i})_i$ (where $\delta_{n,i}$ denotes the Kronecker delta), then clearly

$$\pi_{\Pi,n}(\mathbf{i}; \mathbf{l}) = \pi_n(\mathbf{i}; \mathbf{l}). \quad (2.9)$$

In analogy with the vector $\boldsymbol{\pi}_{k,n}(\mathbf{i}; \mathbf{l}) = (\pi_n(\mathbf{i}; \mathbf{l}))_{\mathbf{l} \in I^k}$ of frequencies of all strings $\mathbf{l} \in I^k$ of length k , we define the vector $\boldsymbol{\pi}_{\Pi,k,n}(\mathbf{i})$ of all Π -average frequencies of strings $\mathbf{l} \in I^k$ of length k by

$$\boldsymbol{\pi}_{\Pi,k,n}(\mathbf{i}) = (\pi_{\Pi,n}(\mathbf{i}; \mathbf{l}))_{\mathbf{l} \in I^k}. \quad (2.10)$$

Analogously to Theorem A, the next result, i.e. Theorem 2.1 below, says that the subsimplex Δ_k^{shift} is also the correct simplex to consider when investigating Π -average frequencies, namely, all accumulation points of the sequence $(\boldsymbol{\pi}_{\Pi,k,n}(\mathbf{i}))_n$ belong to Δ_k^{shift} . We note that the proof of Theorem 2.1 is similar to the proof of [23, Theorem 0]; however, we have decided to include it for completeness.

Theorem 2.1 *Let $\mathbf{i} \in I^{\mathbb{N}}$. Then $\text{acc}_n \boldsymbol{\pi}_{\Pi,k,n}(\mathbf{i}) \subseteq \Delta_k^{\text{shift}}$.*

Proof Let $\mathbf{p} = (p_l)_{l \in I^k} \in \Delta_k$ be an accumulation point of the sequence $(\boldsymbol{\pi}_{\Pi,k,n}(\mathbf{i}))_n$ with respect to $\|\cdot\|_1$, i.e.

$$\liminf_n \|\boldsymbol{\pi}_{\Pi,k,n}(\mathbf{i}) - \mathbf{p}\|_1 = 0. \quad (2.11)$$

By considering all possible ways a string $\mathbf{l} \in \mathbb{N}^{k-1}$ of length $k-1$ can arise as a substring of the string \mathbf{i} it follows that

$$\left| \sum_{l \in I} \pi_n(\mathbf{i}; \mathbf{l}) - \sum_{l \in I} \pi_n(\mathbf{i}; \mathbf{l}) \right| \leq \frac{1}{n}. \quad (2.12)$$

It follows immediately from (2.12) that if $\mathbf{l} \in \mathbb{N}^{k-1}$, then

$$\begin{aligned} \left| \sum_{l \in I} A_{\Pi,n}(\mathbf{i}; \mathbf{l}) - \sum_{l \in I} A_{\Pi,n}(\mathbf{i}; \mathbf{l}) \right| &= \left| \sum_{l \in I} \sum_i \pi_{n,i} \pi_i(\mathbf{i}; \mathbf{l}) - \sum_{l \in I} \sum_i \pi_{n,i} \pi_i(\mathbf{i}; \mathbf{l}) \right| \\ &\leq \sum_i \left| \sum_{l \in I} \pi_{n,i} \pi_i(\mathbf{i}; \mathbf{l}) - \sum_{l \in I} \pi_{n,i} \pi_i(\mathbf{i}; \mathbf{l}) \right| \\ &= \sum_i \pi_{n,i} \left| \sum_{l \in I} \pi_i(\mathbf{i}; \mathbf{l}) - \sum_{l \in I} \pi_i(\mathbf{i}; \mathbf{l}) \right| \leq \frac{1}{n} \sum_i \pi_{n,i} \end{aligned}$$

for all positive integers n , and so

$$\begin{aligned} \left| \sum_{l \in I} p_l - \sum_{l \in I} p_l \right| &\leq \left| \sum_{l \in I} p_l - \sum_{l \in I} A_{\Pi,n}(\mathbf{i}; \mathbf{l}) \right| \\ &\quad + \left| \sum_{l \in I} A_{\Pi,n}(\mathbf{i}; \mathbf{l}) - \sum_{l \in I} A_{\Pi,n}(\mathbf{i}; \mathbf{l}) \right| \\ &\quad + \left| \sum_{l \in I} A_{\Pi,n}(\mathbf{i}; \mathbf{l}) - \sum_{l \in I} p_l \right| \\ &\leq \|\boldsymbol{\pi}_{\Pi,k,n}(\mathbf{i}) - \mathbf{p}\|_1 + \frac{1}{n} \sum_i \pi_{n,i} + \|\boldsymbol{\pi}_{\Pi,k,n}(\mathbf{i}) - \mathbf{p}\|_1 \end{aligned} \quad (2.13)$$

for all n . Since $\frac{1}{n} \sum_i \pi_{n,i} \rightarrow 0$, inequality (2.13) implies that $|\sum_{l \in I} p_l - \sum_{l \in I} p_l| \leq 2 \liminf_n \|\boldsymbol{\pi}_{\Pi,k,n}(\mathbf{i}) - \mathbf{p}\|_1$, and we therefore conclude from (2.11) that $|\sum_{l \in I} p_l -$

$\sum_{l \in I} p_{\mathbf{l}} = 0$, i.e. $\sum_{l \in I} p_{\mathbf{l}} = \sum_{l \in I} p_{\mathbf{l}}$ for all $\mathbf{l} \in \mathbb{N}^{k-1}$, whence $\mathbf{p} = (p_{\mathbf{l}})_{\mathbf{l} \in I^k} \in \Delta_k^{\text{shift}}$. \square

In view of Theorem 2.1 it is natural to attempt to quantify the divergence of the sequence $\pi_{\Pi,k,n}(\mathbf{i}) = (\pi_{\Pi,n}(\mathbf{i}; \mathbf{l}))_{\mathbf{l} \in I^k}$ by considering the extent to which it fills up the subsimplex Δ_k^{shift} . We therefore introduce the following notation. For $\mathbf{i} \in I^{\mathbb{N}}$, we say that the point $\pi(\mathbf{i})$ is extremely non- k - Π -normal if the set of accumulation points of the sequence $(\pi_{\Pi,k,n}(\mathbf{i}))_n$ equals Δ_k^{shift} , and we will denote the set of extremely non- k - Π -normal points by $\mathbb{E}_{\Pi,k}$, i.e.

$$\mathbb{E}_{\Pi,k} = \pi \left\{ \mathbf{i} \in I^{\mathbb{N}} \mid \text{acc}_n \pi_{\Pi,k,n}(\mathbf{i}) = \Delta_k^{\text{shift}} \right\}.$$

We will say that a point is extremely non- Π -normal if it is extremely non- k - Π -normal for all k and we denote the set of extremely non- Π -normal by \mathbb{E}_{Π} , i.e.

$$\mathbb{E}_{\Pi} = \bigcap_k \mathbb{E}_{\Pi,k}.$$

Hence, the points in \mathbb{E}_{Π} are those whose Π -average frequencies of digits diverge in the worst possible way. Our first main result, Theorem 2.3 below, states (perhaps surprisingly) that the set \mathbb{E}_{Π} is extremely big from a topological viewpoint. In the statement of Theorem 2.3 we assume that the function π is closed and it is clearly desirable to have simple conditions guaranteeing that this condition is satisfied. Before stating Theorem 2.3, we immediately give two such conditions. This is the content of the next proposition; in particular, this proposition shows that the map π is closed in the cases considered by Baek, Olsen and Madritsch in Theorem B, and Theorem 2.3 can therefore be applied to the cases considered previously in [2,17,18,23].

Proposition 2.2 *Let $(X, (S_i)_{i \in I})$ be an IIFS and assume that one of the following two conditions (a) or (b) is satisfied:*

- (a) *The set I is finite and the list $(X, (S_i)_{i \in I})$ satisfies the SOSC;*
- (b) *The list $(X, (S_i)_{i \in I})$ is the continued fraction IIFS, i.e. $I = \mathbb{N}$, $X = [0, 1]$ and $S_i : [0, 1] \rightarrow [0, 1]$ is defined by $S_i(x) = \frac{1}{x+i}$ for all i .*

Then the function π is closed.

Proof If I is finite, then $K = \pi(I^{\mathbb{N}})$ is compact, and since $\pi : I^{\mathbb{N}} \rightarrow K$ is continuous, we therefore conclude that π maps closed sets to compact sets and hence to closed sets.

On the other hand, if the list $(X, (S_i)_{i \in I})$ is the continued fraction IIFS, then it is well-known that the map $\pi : I^{\mathbb{N}} \rightarrow K$ is a homeomorphism and hence closed. \square

We can now state Theorem 2.3.

Theorem 2.3 *Let $(X, (S_i)_{i \in I})$ be an IIFS. Assume that $(X, (S_i)_{i \in I})$ satisfies the SOSC and that the map π is closed. Then the set \mathbb{E}_{Π} is comeager in K .*

The proof of Theorem 2.3 is given in Sects. 6 and 7. Several applications of Theorem 2.3 will be presented in Sects. 3–5: in Sect. 3 we use Theorem 2.3 to investigate higher order Hölder and Cesaro averages, in Sect. 4 we use Theorem 2.3 to investigate logarithmic Reisz averages, and in Sect. 5 we apply Theorem 2.3 to study the distribution of continued fraction digits and Lüroth expansion digits.

We also find the packing dimension of the set \mathbb{E}_Π . Indeed, the next result shows that if the assumptions in Theorem 2.3 and an additional mild condition on the maps S_i are satisfied, then the packing dimension of \mathbb{E}_Π is maximal, i.e. equal to the packing dimension of the ambient set K . This result may be viewed as a further manifestation of the fact that the set \mathbb{E}_Π is “big”. Below we denote the packing dimension by $\dim_P(F)$; the reader is referred to [8] for the definition of the packing dimension.

Theorem 2.4 *Let $(X, (S_i)_{i \in I})$ be an IIFS. Assume that $(X, (S_i)_{i \in I})$ satisfies the SOSC and that the map π is closed. Assume further that the maps S_i are bi-Lipschitz for all i . Then*

$$\dim_P(\mathbb{E}_\Pi) = \dim_P(K).$$

The proof of Theorem 2.4 is given in Sect. 8.

We have not considered the problem of computing the Hausdorff dimension of the set \mathbb{E}_Π . However, it follows from [24,25] that the Hausdorff dimension of the set of numbers whose N -adic-expansion satisfies a similar condition of extreme non-normality equals 0, and it follows from [16] that the Hausdorff dimension of the set of numbers whose continued fraction expansion satisfies a similar condition of extreme non-normality equals 0. We therefore make the following conjecture. Below we denote the Hausdorff dimension by \dim_H ; the reader is referred to [8] for the definition of the Hausdorff dimension.

Conjecture 2.5 *Let $(X, (S_i)_{i \in I})$ be an IIFS. Assume that $(X, (S_i)_{i \in I})$ satisfies the SOSC and that the map π is closed. Assume further that the maps S_i are bi-Lipschitz for all i . Then*

$$\dim_H(\mathbb{E}_\Pi) = 0.$$

In fact, for each positive integer k , we have

$$\dim_H(\mathbb{E}_{\Pi,k}) = 0.$$

2.5 Points of maximal oscillation

Our results can be applied to the study of points of maximal oscillation. Recently such points have received considerable interest in the literature and it has been proved that in many cases typical points have maximal oscillation. A point $\pi(\mathbf{i})$ is said to have maximal oscillation if the frequencies $\pi_n(\mathbf{i}; i)$ of all digits i oscillate as much as possible as $n \rightarrow \infty$, more precisely, a point $\pi(\mathbf{i})$ is said to have maximal oscillation

if $\underline{\pi}(\mathbf{i}; i) = 0$ and $\overline{\pi}(\mathbf{i}; i) = 1$ for all digits i ; we let \mathbb{M} denote the set of points of maximal oscillation, i.e.

$$\mathbb{M} = \pi \left\{ \mathbf{i} \in I^{\mathbb{N}} \mid \underline{\pi}(\mathbf{i}; i) = 0 \text{ and } \overline{\pi}(\mathbf{i}; i) = 1 \text{ for all } i \in I \right\}. \quad (2.14)$$

Šalat [30] proved that if $(X, (S_i)_{i \in I})$ is the continued fraction IIFS from Sect. 5.1 (and, hence, the digits \mathbf{i} are the continued fraction digits of $x = \pi(\mathbf{i})$), then \mathbb{M} is comeagre, and Madritsch [18] proved analogous results for the frequencies of strings of continued fraction digits. We will now show that Theorem 2.3 provides substantial extensions of these and many other results.

It is natural to consider points for which the average of the frequencies of strings of digits have maximal oscillation. For this reason we introduce the following notation. Fix $k \in \mathbb{N}$ and $\mathbf{l} \in I^k$. Define the projection $P_{k, \mathbf{l}} : \mathbb{R}^{I^k} \rightarrow \mathbb{R}$ onto the \mathbf{l} 'th coordinate by

$$P_{k, \mathbf{l}}((x_{\mathbf{i}})_{\mathbf{i} \in I^k}) = x_{\mathbf{l}},$$

and write

$$\begin{aligned} u_{k, \mathbf{l}} &= \inf P_{k, \mathbf{l}}(\Delta_k^{\text{shift}}), \\ v_{k, \mathbf{l}} &= \sup P_{k, \mathbf{l}}(\Delta_k^{\text{shift}}). \end{aligned}$$

We now define the set of points of maximal Π - k -oscillation by

$$\mathbb{M}_{\Pi, k} = \pi \left\{ \mathbf{i} \in I^{\mathbb{N}} \mid \underline{\pi}_{\Pi}(\mathbf{i}; \mathbf{l}) = u_{k, \mathbf{l}} \text{ and } \overline{\pi}_{\Pi}(\mathbf{i}; \mathbf{l}) = v_{k, \mathbf{l}} \text{ for all } \mathbf{l} \in I^k \right\},$$

and we define the set of points of maximal Π -oscillation by

$$\mathbb{M}_{\Pi} = \bigcap_k \mathbb{M}_{\Pi, k}.$$

Remark We note that the notion of points of maximal Π - k -oscillation extends the classical notion of points of maximal oscillation in (2.14); indeed, if we let Π denote the trivial averaging system defined by $\Pi = (\Pi_n)_n$ where $\Pi_n = (\delta_{n, i})_i$, then clearly $\mathbb{M}_{\Pi, 1} = \mathbb{M}$.

Remark Since $\Delta_1^{\text{shift}} = \Delta_1$, we have

$$u_{1, \mathbf{l}} = 0, \quad v_{1, \mathbf{l}} = 1.$$

However, for $k > 1$, it may happen that $v_{k, \mathbf{l}}$ is strictly less than 1. For example, if we let $I = \{0, 1\}$, $k = 2$ and $\mathbf{l} = 10$, then

$$v_{k, \mathbf{l}} < 1.$$

Indeed, otherwise $\sup P_{k, \mathbf{l}}(\Delta_k^{\text{shift}}) = v_{k, \mathbf{l}} = 1$, and since $P_{k, \mathbf{l}}(\Delta_k^{\text{shift}})$ is compact (because Δ_k^{shift} is a closed and bounded subset of $\mathbb{R}^{I^k} = \mathbb{R}^{\{0,1\}^2} = \mathbb{R}^4$ (and hence compact) and $P_{k, \mathbf{l}}$ is continuous), we now conclude that there is a probability vector $\mathbf{p} = (p_i)_{i \in I^k} \in \Delta_k^{\text{shift}}$ with $P_{k, \mathbf{l}}(\mathbf{p}) = 1$. However, this clearly implies that $p_{\mathbf{l}} = 1$ and $p_i = 0$ for $\mathbf{i} \neq \mathbf{l}$, and so $\sum_{i=0,1} p_{i0} = 1$ and $\sum_{i=0,1} p_{0i} = 0$, whence $\mathbf{p} \notin \Delta_k^{\text{shift}}$, contradicting the fact that $\mathbf{p} \in \Delta_k^{\text{shift}}$.

The next result says that typical points have maximal oscillation.

Theorem 2.6 *Let $(X, (S_i)_{i \in I})$ be an IIFS. Assume that $(X, (S_i)_{i \in I})$ satisfies the SOSC and that the map π is closed. Then the set \mathbb{M}_{Π} of points of maximal Π -oscillation is comeagre. In particular, the set \mathbb{M} (in (2.14)) of points of maximal oscillation is comeagre.*

Proof For $\mathbf{l} \in I^k$, write

$$\mathbb{M}_{\Pi, k, \mathbf{l}} = \pi \left\{ \mathbf{i} \in I^{\mathbb{N}} \mid \pi_{\Pi}(\mathbf{i}; \mathbf{l}) = v_{k, \mathbf{l}} \text{ and } \bar{\pi}_{\Pi}(\mathbf{i}; \mathbf{l}) = u_{k, \mathbf{l}} \right\}.$$

Since $\mathbb{E}_{\Pi, k}$ is comeagre for $k \in \mathbb{N}$ (by Theorem 2.3) and $\mathbb{M}_{\Pi, k} = \bigcap_{\mathbf{l} \in I^k} \mathbb{M}_{\Pi, k, \mathbf{l}}$, it clearly suffices to show that $\mathbb{E}_{\Pi, k} \subseteq \mathbb{M}_{\Pi, k, \mathbf{l}}$ for all $k \in \mathbb{N}$ and all $\mathbf{l} \in I^k$. To prove this, we fix $k \in \mathbb{N}$ and $\mathbf{l} \in I^k$, and let $\mathbf{i} \in \mathbb{E}_{\Pi, k}$. It follows from the definitions of $u_{k, \mathbf{l}}$ and $v_{k, \mathbf{l}}$ that we can choose sequences $(\mathbf{u}_j)_j$ and $(\mathbf{v}_j)_j$ in Δ_k^{shift} such that $P_{k, \mathbf{i}}(\mathbf{u}_j) \rightarrow u_{k, \mathbf{l}}$ and $P_{k, \mathbf{i}}(\mathbf{v}_j) \rightarrow v_{k, \mathbf{l}}$. Also, since $\mathbf{i} \in \mathbb{E}_{\Pi, k}$, we can choose a strictly increasing sequence $(n_j)_j$ of positive integers with $\|\pi_{\Pi, k, n_{2j}}(\mathbf{i}) - \mathbf{u}_j\|_1 \rightarrow 0$ and $\|\pi_{\Pi, k, n_{2j+1}}(\mathbf{i}) - \mathbf{v}_j\|_1 \rightarrow 0$. This clearly implies that $\pi_{\Pi, n_{2j}}(\mathbf{i}; \mathbf{l}) - P_{k, \mathbf{l}}(\mathbf{u}_j) \rightarrow 0$ and $\pi_{\Pi, n_{2j+1}}(\mathbf{i}; \mathbf{l}) - P_{k, \mathbf{l}}(\mathbf{v}_j) \rightarrow 0$, and since $P_{k, \mathbf{i}}(\mathbf{u}_j) \rightarrow u_{k, \mathbf{l}}$ and $P_{k, \mathbf{i}}(\mathbf{v}_j) \rightarrow v_{k, \mathbf{l}}$, we therefore conclude that $\pi_{\Pi, n_{2j}}(\mathbf{i}; \mathbf{l}) \rightarrow u_{k, \mathbf{l}}$ and $\pi_{\Pi, n_{2j+1}}(\mathbf{i}; \mathbf{l}) \rightarrow v_{k, \mathbf{l}}$, whence $\mathbf{i} \in \mathbb{M}_{\Pi, k, \mathbf{l}}$. \square

2.6 Comparing with normal points

Our main results show that the limiting frequencies of digits of a typical point in K fail to exist in a very spectacular way. This is often in very sharp contrast to the behaviour of the limiting frequencies of digits of almost all points in K with respect to judiciously chosen ergodic measures, and we believe that it is instructive to provide some context by briefly explaining this. For $i \in I$, we define the cylinder $[i]$ generated by i by

$$[i] = \left\{ \mathbf{j} = j_1 j_2 \dots \in I^{\mathbb{N}} \mid j_1 = i \right\}.$$

In analogy with the usual definition of normal numbers (see, for example, [14]), we now make the following definition.

Definition Normal Let μ be a probability measure on $I^{\mathbb{N}}$. Write

$$N_{\mu} = \left\{ \mathbf{i} \in I^{\mathbb{N}} \mid \pi_n(\mathbf{i}; i) \rightarrow \mu([i]) \text{ for all } i \in I \right\}.$$

We now define the set of μ -normal points in K by

$$\mathbb{N}_\mu = \pi(N_\mu) = \pi \left\{ \mathbf{i} \in I^{\mathbb{N}} \mid \pi_n(\mathbf{i}; i) \rightarrow \mu([i]) \text{ for all } i \in I \right\}.$$

Remark It is clear that the above definition of normality extends the classical notion of normal numbers. Indeed, fix a positive integer N . Write $I = \{0, \dots, N-1\}$ and for $i \in I$, define $S_i : [0, 1] \rightarrow [0, 1]$ by $S_i(x) = \frac{x+i}{N}$. Let μ denote the Bernoulli measure on $I^{\mathbb{N}}$ such that $\mu([i]) = \frac{1}{N}$ for all i , and for $x \in [0, 1]$, let $x = \sum_n \frac{d_n(x)}{N^n}$ denote the unique non-terminating N -adic expansion of x . With these definitions, it is not difficult to see that

$$\mathbb{N}_\mu = \left\{ x \in [0, 1] \mid \frac{|\{1 \leq j \leq n \mid d_j(x) = i\}|}{n} \rightarrow \frac{1}{N} \text{ for all } i \right\},$$

i.e. \mathbb{N}_μ equals the set of numbers that are normal to base N .

As remarked above, our main results show that the limiting frequencies of digits of a typical point in K fail to exist in a very spectacular way. We also suggested that this often is in very sharp contrast to the behaviour of the limiting frequencies of digits of almost all points in K with respect to judiciously chosen ergodic measures. We can now make this statement precise. Namely, in many important and natural cases μ -almost all points in K are μ -normal. Indeed, if μ is a shift-invariant and ergodic probability measure on $I^{\mathbb{N}}$ and we write $\hat{\mu} = \mu \circ \pi^{-1}$ for the image measure of μ on K , then it follows immediately from the ergodic theorem that $\mu(N_\mu) = 1$, and since $\pi^{-1}(K \setminus \mathbb{N}_\mu) = \pi^{-1}(K \setminus \pi(N_\mu)) \subseteq I^{\mathbb{N}} \setminus N_\mu$, we therefore conclude that $\hat{\mu}(K \setminus \mathbb{N}_\mu) = \mu(\pi^{-1}(K \setminus \mathbb{N}_\mu)) \leq \mu(I^{\mathbb{N}} \setminus N_\mu) = 0$, whence

$$\hat{\mu}(\mathbb{N}_\mu) = 1,$$

i.e. $\hat{\mu}$ -almost all points in K are μ -normal. Hence, often the limiting frequencies of digits of almost all points in K exist and equal the “correct” average value.

3 Hölder and Cesaro averages of frequencies of digits in IIFS's

Two of the most important and commonly used averages are the Hölder averages and the Cesaro averages. We will now define these averages and apply them to the sequence $\boldsymbol{\pi}(\mathbf{i}; \mathbf{l}) = (\pi_n(\mathbf{i}; \mathbf{l}))_n$ of frequencies of the string \mathbf{l} in string \mathbf{i} . We start by recalling the definitions of the Hölder and Cesaro averages. For a sequence $\mathbf{x} = (x_n)_n$ of real numbers and an integer m with $m \geq 1$, we define the sequence $(H_n^m(\mathbf{x}))_n$ inductively as follows

$$H_n^0(\mathbf{x}) = x_n, \quad H_n^m(\mathbf{x}) = \frac{H_1^{m-1}(\mathbf{x}) + \dots + H_n^{m-1}(\mathbf{x})}{n}.$$

The lower and upper m 'th order Hölder averages of \mathbf{x} are now defined by

$$\underline{H}^m(\mathbf{x}) = \liminf_n H_n^m(\mathbf{x}),$$

$$\overline{H}^m(\mathbf{x}) = \limsup_n H_n^m(\mathbf{x}).$$

The Cesaro averages are defined as follows. For a sequence $\mathbf{x} = (x_n)_n$ of real numbers and an integer m with $m \geq 1$, we define the sequence $(C_n^m(\mathbf{x}))_n$ inductively as follows

$$C_n^0(\mathbf{x}) = x_n, \quad C_n^m(\mathbf{x}) = C_1^{m-1}(\mathbf{x}) + \dots + C_n^{m-1}(\mathbf{x}).$$

The lower and upper m 'th order Cesaro averages of \mathbf{x} are now defined by

$$\underline{C}^m(\mathbf{x}) = \liminf_n \frac{1}{\binom{n+m}{m}} C_n^m(\mathbf{x}),$$

$$\overline{C}^m(\mathbf{x}) = \limsup_n \frac{1}{\binom{n+m}{m}} C_n^m(\mathbf{x}).$$

It is well-known that the higher order Hölder and Cesaro averages form a double infinite hierarchy between the lower limit of the sequence \mathbf{x} and the upper limit of the sequence \mathbf{x} in (at least) countably infinite many levels, namely, we have the following inequalities

$$\liminf_n x_n = \underline{H}^0(\mathbf{x}) \leq \underline{H}^1(\mathbf{x}) \leq \underline{H}^2(\mathbf{x}) \leq \dots \leq \overline{H}^2(\mathbf{x}) \leq \overline{H}^1(\mathbf{x}) \leq \overline{H}^0(\mathbf{x}) = \limsup_n x_n,$$

$$\liminf_n x_n = \underline{C}^0(\mathbf{x}) \leq \underline{C}^1(\mathbf{x}) \leq \underline{C}^2(\mathbf{x}) \leq \dots \leq \overline{C}^2(\mathbf{x}) \leq \overline{C}^1(\mathbf{x}) \leq \overline{C}^0(\mathbf{x}) = \limsup_n x_n. \quad (3.1)$$

It is also well-known that the Hölder and Cesaro averages are averaging systems in the sense of the definition in Sect. 2.4. Indeed, if we for a positive integer n , define the averaging system $\Pi_{\mathbf{H}}^m = (\Pi_{\mathbf{H},n}^m)_n$ by $\Pi_{\mathbf{H},n}^m = (\pi_{\mathbf{H},n,i}^m)_i$ where

$$\pi_{\mathbf{H},n,i}^m = \begin{cases} \sum_{j=i}^n (-1)^{j+i} \frac{1}{(j+1)^m} \binom{n}{j} \binom{j}{i} & \text{for } i = 1, \dots, n; \\ 0 & \text{for } i > n, \end{cases}$$

then $H_n^m(\mathbf{x}) = A_{\Pi_{\mathbf{H},n}^m}(\mathbf{x})$, see, for example, [9, Chap XI, §11.4], and so $\underline{H}^m(\mathbf{x}) = \underline{A}_{\Pi_{\mathbf{H}}^m}(\mathbf{x})$ and $\overline{H}^m(\mathbf{x}) = \overline{A}_{\Pi_{\mathbf{H}}^m}(\mathbf{x})$. Similarly, if we for a positive integer m , define the averaging system $\Pi_{\mathbf{C}}^m = (\Pi_{\mathbf{C},n}^m)_n$ by $\Pi_{\mathbf{C},n}^m = (\pi_{\mathbf{C},n,i}^m)_i$ where

$$\pi_{\mathbf{C},n,i}^m = \begin{cases} \frac{1}{\binom{n+m}{m}} \binom{n-i+m-1}{m-1} & \text{for } i = 1, \dots, n; \\ 0 & \text{for } i > n, \end{cases}$$

then $\frac{1}{\binom{n+m}{m}} C_n^m(\mathbf{x}) = A_{\Pi_{\mathbf{C},n}^m}(\mathbf{x})$, see, for example, [9, Chap V, §5.4 or Chap XI, §11.4] or [13], and so $\underline{C}^m(\mathbf{x}) = \underline{A}_{\Pi_{\mathbf{C}}^m}(\mathbf{x})$ and $\overline{C}^m(\mathbf{x}) = \overline{A}_{\Pi_{\mathbf{C}}^m}(\mathbf{x})$.

Using Hölder and Cesaro averages we can now introduce average Hölder and Cesaro frequencies by applying the definitions of the Hölder and Cesaro averages to the sequence $\pi(\mathbf{i}; \mathbf{I}) = (\pi_n(\mathbf{i}; \mathbf{I}))_n$ of frequencies of the string \mathbf{I} in string \mathbf{i} . This is the content of the next definition.

Definition *Higher order Hölder and Cesaro averages frequencies of digits* Let $\mathbf{i} \in I^{\mathbb{N}}$. For a finite string $\mathbf{I} \in I^k$, we define the lower and upper n 'th order Hölder average frequency of the string \mathbf{I} among the digits of the string \mathbf{i} by

$$\begin{aligned}\underline{\pi}_H^m(\mathbf{i}; \mathbf{I}) &= \underline{H}^m(\pi(\mathbf{i}; \mathbf{I})) = \liminf_n H_n^m(\pi(\mathbf{i}; \mathbf{I})), \\ \overline{\pi}_H^m(\mathbf{i}; \mathbf{I}) &= \overline{H}^m(\pi(\mathbf{i}; \mathbf{I})) = \limsup_n H_n^m(\pi(\mathbf{i}; \mathbf{I})).\end{aligned}$$

Similarly, we define the lower and upper n 'th order Cesaro average frequency of the string \mathbf{I} among the digits of the string \mathbf{i} by

$$\begin{aligned}\underline{\pi}_C^m(\mathbf{i}; \mathbf{I}) &= \underline{C}^m(\pi(\mathbf{i}; \mathbf{I})) = \liminf_n \frac{1}{\binom{n+m}{m}} C_n^m(\pi(\mathbf{i}; \mathbf{I})), \\ \overline{\pi}_C^m(\mathbf{i}; \mathbf{I}) &= \overline{C}^m(\pi(\mathbf{i}; \mathbf{I})) = \limsup_n \frac{1}{\binom{n+m}{m}} C_n^m(\pi(\mathbf{i}; \mathbf{I})).\end{aligned}$$

The higher order average Hölder and Cesaro frequencies form a double infinite hierarchy between the lower frequency and the upper frequency in (at least) countably infinite many levels, namely, we have (using (3.1))

$$\begin{aligned}\underline{\pi}(\mathbf{i}; \mathbf{I}) &= \underline{\pi}_H^0(\mathbf{i}; \mathbf{I}) \leq \underline{\pi}_H^1(\mathbf{i}; \mathbf{I}) \leq \underline{\pi}_H^2(\mathbf{i}; \mathbf{I}) \leq \dots \leq \overline{\pi}_H^2(\mathbf{i}; \mathbf{I}) \leq \overline{\pi}_H^1(\mathbf{i}; \mathbf{I}) \leq \overline{\pi}_H^0(\mathbf{i}; \mathbf{I}) = \overline{\pi}(\mathbf{i}; \mathbf{I}), \\ \underline{\pi}(\mathbf{i}; \mathbf{I}) &= \underline{\pi}_C^0(\mathbf{i}; \mathbf{I}) \leq \underline{\pi}_C^1(\mathbf{i}; \mathbf{I}) \leq \underline{\pi}_C^2(\mathbf{i}; \mathbf{I}) \leq \dots \leq \overline{\pi}_C^2(\mathbf{i}; \mathbf{I}) \leq \overline{\pi}_C^1(\mathbf{i}; \mathbf{I}) \leq \overline{\pi}_C^0(\mathbf{i}; \mathbf{I}) = \overline{\pi}(\mathbf{i}; \mathbf{I}).\end{aligned}\quad (3.2)$$

As an application of Theorem 2.3, we will now show that the behaviour of the digits of a typical point in K is so irregular that not even the hierarchies in (3.2) formed by taking Hölder and Cesaro averages of *all* orders are sufficiently powerful to “smoothen out” the behaviour of the frequencies of the digits. Let

$$\pi_{H,k,n}^m(\mathbf{i}) = (H_n^m(\pi(\mathbf{i}; \mathbf{I})))_{\mathbf{I} \in I^k}, \quad (3.3)$$

i.e. $\pi_{H,k,n}^m(\mathbf{i})$ denotes the vector of m 'th order Hölder-average frequencies $H_n^m(\pi(\mathbf{i}; \mathbf{I}))$ of all strings $\mathbf{I} \in I^k$ of length k . Similarly, we let

$$\pi_{C,k,n}^m(\mathbf{i}) = \left(\frac{1}{\binom{n+m}{m}} C_n^m(\pi(\mathbf{i}; \mathbf{I})) \right)_{\mathbf{I} \in I^k}, \quad (3.4)$$

i.e. $\pi_{C,k,n}^m(\mathbf{i})$ denotes the vector of m 'th order Cesaro-average frequencies $C_n^m(\pi(\mathbf{i}; \mathbf{I}))$ of all strings $\mathbf{I} \in I^k$ of length k .

For $\mathbf{i} \in I^{\mathbb{N}}$, we say that the point $\pi(\mathbf{i})$ is extremely non-Hölder-normal if the set of accumulation points of the sequence $(\pi_{H,k,n}^m(\mathbf{i}))_n$ (with respect to $\|\cdot\|_1$) equals Δ_k^{shift}

for all m and all k , and we will denote the set of extremely non-Hölder-normal points by \mathbb{E}_H , i.e.

$$\mathbb{E}_H = \bigcap_{m,k} \pi \left\{ \mathbf{i} \in I^k \mid \operatorname{acc}_n \pi_{H,k,n}^m(\mathbf{i}) = \Delta_k^{\text{shift}} \right\}.$$

Similarly, we define the set of extremely non-Cesaro-normal points, denoted by \mathbb{E}_C , by

$$\mathbb{E}_C = \bigcap_{m,k} \pi \left\{ \mathbf{i} \in I^k \mid \operatorname{acc}_n \pi_{C,k,n}^m(\mathbf{i}) = \Delta_k^{\text{shift}} \right\}.$$

Hence, the points in \mathbb{E}_H and \mathbb{E}_C are as far away from being normal that *all* higher order Hölder and Cesaro averages of frequencies of strings of all lengths diverges in in the worst possible way, i.e. the points in \mathbb{E}_H and \mathbb{E}_C are so irregular that not even the hierarchies in (3.2) formed by taking Hölder and Cesaro averages of *all* orders are sufficiently powerful to “smoothen out” the behaviour of the frequencies of their digits. The theorem below shows (perhaps surprisingly), that the sets \mathbb{E}_H and \mathbb{E}_C are extremely big from a topological viewpoint.

Theorem 3.1 *Let $(X, (S_i)_{i \in I})$ be an IIFS. Assume that $(X, (S_i)_{i \in I})$ satisfies the SOSC and that the map π is closed. Then the sets \mathbb{E}_H and \mathbb{E}_C are comeager in K .*

Proof Since $H_n^m(\mathbf{x}) = A_{\Pi_{H,n}^m}(\mathbf{x})$ and $C_n^m(\mathbf{x}) = A_{\Pi_{C,n}^m}(\mathbf{x})$ for all real valued sequences \mathbf{x} and an easy calculation shows that $\frac{1}{n} \sum_i \pi_{H,n,i}^m \rightarrow 0$ as $n \rightarrow \infty$ and $\frac{1}{n} \sum_i \pi_{C,n,i}^m \rightarrow 0$ as $n \rightarrow \infty$, the statement in Theorem 3.1 follows immediately from Theorem 2.3. \square

4 Logarithmic Riesz averages

A further important average method, that is even more powerful than *all* higher order Cesaro averages, is the logarithmic Riesz average. This method was introduced by Marcel Riesz [28] in the early 1900’s and is of particular importance in the study of convergence of Dirichlet series, see, for example, Hardy and Riesz’s classical text on Dirichlet series [10]. We now recall the definition of the logarithmic Riesz average and apply it to the sequence $\pi(\mathbf{i}; \mathbf{l}) = (\pi_n(\mathbf{i}; \mathbf{l}))_n$ of frequencies. For a sequence $\mathbf{x} = (x_n)_n$ of real numbers, define the sequence $(R_n(\mathbf{x}))_n$ by

$$R_n(\mathbf{x}) = \frac{1}{\log(n+1)} \left(\frac{x_1}{1} + \frac{x_2}{2} + \cdots + \frac{x_n}{n} \right).$$

The lower and upper logarithmic Riesz averages of \mathbf{x} are now defined by

$$\begin{aligned} \underline{R}(\mathbf{x}) &= \liminf_n R_n(\mathbf{x}), \\ \overline{R}(\mathbf{x}) &= \limsup_n R_n(\mathbf{x}), \end{aligned}$$

see, for example, [9, Chap IV, §4.16]. It is well-known that the logarithmic Riesz average is more powerful than *all* higher order Cesaro averages in the following sense: if the m 'th order Cesaro average of the sequence \mathbf{x} exists and equal s , then the logarithmic Riesz average of \mathbf{x} also exists and equals s , and there are sequences for which none of the m 'th order Cesaro averages exist but the logarithmic Riesz average exists, see [9, Chap V, §5.13.(b)]. More precisely, the following inequalities hold

$$\underline{C}^m(\mathbf{x}) \leq \underline{R}(\mathbf{x}) \leq \overline{R}(\mathbf{x}) \leq \overline{C}^m(\mathbf{x}) \quad (4.1)$$

for all $m \geq 0$; while the inequalities in (4.1) undoubtedly are well-known we have been unable to find a proof in the literature and for the convenience of the reader we have decided to provide a short and direct proof of (4.1) in "Appendix A". It is also clear that the logarithmic Riesz averages are averaging systems in the sense of the definition in Sect. 2.4. Indeed, if we for a positive integer n , define the averaging system $\Pi_R = (\Pi_{R,n})_n$ by $\Pi_{R,n} = (\pi_{R,n,i})_i$ where

$$\pi_{R,n,i} = \begin{cases} \frac{1}{\log(n+1)} \frac{1}{i} & \text{for } i = 1, \dots, n; \\ 0 & \text{for } i > n, \end{cases}$$

then clearly $R_n(\mathbf{x}) = A_{\Pi_{R,n}}(\mathbf{x})$, and so $\underline{R}(\mathbf{x}) = \underline{A}_{\Pi_R}(\mathbf{x})$ and $\overline{R}(\mathbf{x}) = \overline{A}_{\Pi_R}(\mathbf{x})$.

Using logarithmic Riesz averages we can now introduce logarithmic Riesz average frequencies by applying the definitions of the logarithmic Riesz averages to the sequence $\boldsymbol{\pi}(\mathbf{i}; \mathbf{l}) = (\pi_n(\mathbf{i}; \mathbf{l}))_n$ of frequencies of the string \mathbf{l} in string \mathbf{i} . This is the content of the next definition.

Definition *Logarithmic Riesz averages frequencies of digits* Let $\mathbf{i} \in I^{\mathbb{N}}$. For a finite string $\mathbf{l} \in I^k$, we define the lower and upper logarithmic Riesz average frequency of the string \mathbf{l} among the digits of the string \mathbf{i} by

$$\begin{aligned} \underline{\pi}_R(\mathbf{i}; \mathbf{l}) &= \underline{R}(\boldsymbol{\pi}(\mathbf{i}; \mathbf{l})) = \liminf_n R_n(\boldsymbol{\pi}(\mathbf{i}; \mathbf{l})), \\ \overline{\pi}_R(\mathbf{i}; \mathbf{l}) &= \overline{R}(\boldsymbol{\pi}(\mathbf{i}; \mathbf{l})) = \limsup_n R_n(\boldsymbol{\pi}(\mathbf{i}; \mathbf{l})). \end{aligned}$$

The logarithmic Riesz average frequencies are more powerful than *all* higher order average Cesaro frequencies, namely, we have (using (4.1))

$$\underline{\pi}_C^m(\mathbf{i}; \mathbf{l}) \leq \underline{\pi}_R(\mathbf{i}; \mathbf{l}) \leq \overline{\pi}_R(\mathbf{i}; \mathbf{l}) \leq \overline{\pi}_C^m(\mathbf{i}; \mathbf{l})$$

for all m .

As an application of Theorem 2.3, we will now show that the behaviour of the digits of a typical point in K is so irregular that not even taking logarithmic Riesz averages is sufficiently powerful to "smoothen out" the behaviour of the frequencies of the digits. Let

$$\boldsymbol{\pi}_{R,k,n}(\mathbf{i}) = (R_n(\boldsymbol{\pi}(\mathbf{i}; \mathbf{l})))_{\mathbf{l} \in I^k},$$

i.e. $\pi_{R,k,n}(\mathbf{i})$ denotes the vector of the logarithmic Riesz average frequencies $R_n(\pi(\mathbf{i}; \mathbf{l}))$ of all strings $\mathbf{l} \in I^k$ of length k . For $\mathbf{i} \in I^{\mathbb{N}}$, we say that the point $\pi(\mathbf{i})$ is extremely non-logarithmic-Riesz-normal if the set of accumulation points of the sequence $(\pi_{R,k,n}(\mathbf{i}))_n$ (with respect to $\|\cdot\|_1$) equals Δ_k^{shift} for all k , and we will denote the set of extremely non-logarithmic-Riesz-normal points by \mathbb{E}_R , i.e.

$$\mathbb{E}_R = \bigcap_k \pi \left\{ \mathbf{i} \in I^k \mid \text{acc}_n \pi_{R,k,n}(\mathbf{i}) = \Delta_k^{\text{shift}} \right\}.$$

Hence, the points in \mathbb{E}_R are so far away from being normal that even forming the powerful logarithmic Riesz average is not sufficiently powerful to “smoothen out” the behaviour of the frequencies of their digits. The theorem below shows that the set \mathbb{E}_R is extremely big from a topological viewpoint.

Theorem 4.1 *Let $(X, (S_i)_{i \in I})$ be an IIFS. Assume that $(X, (S_i)_{i \in I})$ satisfies the SOSC and that the map π is closed. Then the set \mathbb{E}_R is comeager in K .*

Proof Since $R_n(\mathbf{x}) = A_{\Pi_{R,n}}(\mathbf{x})$ for all real valued sequences \mathbf{x} and an easy calculation shows that $\frac{1}{n} \sum_i \pi_{R,n,i} \rightarrow 0$ as $n \rightarrow \infty$, the statement in Theorem 4.1 follows immediately from Theorem 2.3. \square

5 Examples

5.1 Continued fraction expansion

Recall, that \mathbb{N} denotes the set of all strictly positive integers, i.e. $\mathbb{N} = \{1, 2, \dots\}$; in particular, $0 \notin \mathbb{N}$. Let \mathbb{P} denote the irrational numbers in the closed unit interval, i.e. $\mathbb{P} = [0, 1] \setminus \mathbb{Q}$. For $x \in \mathbb{P}$, let

$$x = \frac{1}{a_1(x) + \frac{1}{a_2(x) + \dots}}, \quad (5.1)$$

where $a_n(x) \in \mathbb{N}$ for all n , denote the simple continued fraction expansion of x , and let $\mathbf{a}(x)$ denote the string of the continued fraction digits of x , i.e.

$$\mathbf{a}(x) = a_1(x) a_2(x) \dots \in \mathbb{N}^{\mathbb{N}}.$$

In order to analyse the frequencies of the continued fraction digits of x , we define the IIFS $(X, (S_i)_{i \in I})$ as follows. Put $X = [0, 1]$, $I = \mathbb{N}$ and for $i \in \mathbb{N}$, define $S_i : [0, 1] \rightarrow [0, 1]$ by

$$S_i(x) = \frac{1}{x + i}. \quad (5.2)$$

For this choice of X and S_i it is well-known that $\pi(\mathbb{N}^{\mathbb{N}}) = \mathbb{P}$ and that π is a homeomorphism between $\mathbb{N}^{\mathbb{N}}$ and \mathbb{P} . It is also well-known that $\pi^{-1}(x) = \mathbf{a}(x)$ for all $x \in \mathbb{P}$;

it is, of course, this identity that allow us to apply Theorem 2.3 to the IIFS in (5.2) in order to obtain results about the continued fraction digits of a typical number $x \in \mathbb{P}$.

Recall that if n is a positive integer and $i \in \mathbb{N}$ is a digit, then $\pi_n(\mathbf{i}; i) = \frac{|\{1 \leq j \leq n \mid i_j = i\}|}{n}$ denotes the frequency of the digit i among the first n digits of the string $\mathbf{i} = i_1 i_2 \dots$. In particular, $\pi_n(\mathbf{a}(x); i)$ denotes the frequency of the digit i among the first n digits of the continued fraction digits of x . A classical result due to Lévy [15] says that for Lebesgue almost all $x \in \mathbb{P}$ we have

$$\pi_n(\mathbf{a}(x); i) \rightarrow \frac{1}{\log 2} \log \frac{(i+1)^2}{i(i+2)} \quad (5.3)$$

for all $i \in \mathbb{N}$; the reader is referred to the textbook [4, p 45] for a contemporary proof of this based on the ergodic theorem. In analogy with normal numbers, we will say that a number $x \in \mathbb{P}$ is continued fraction normal (c-f-normal) if it satisfies (5.3). Hence, using this terminology, Lévy's result says that Lebesgue almost all $x \in \mathbb{P}$ are c-f-normal. As an application of Theorem 2.3 to the IIFS in (5.2), we will now show that the behaviour of the continued fraction digits of a typical point x in \mathbb{P} is so irregular that *no* averaging system (regardless of how powerful it is) is powerful enough to “smoothen out” the behaviour of the frequencies of the digits; more precisely, for *all* averaging systems Π (regardless of how powerful they are), the set of accumulation points of the sequence $(\pi_{\Pi, k, n}(\mathbf{a}(x)))_n$ of Π -average frequencies of a typical number x is as big as possible. This is the content of the next theorem.

Theorem 5.1 *Let Π be an averaging system and write*

$$\mathbb{E}_{\Pi}^{\text{cf}} = \bigcap_k \left\{ x \in \mathbb{P} \mid \text{acc}_n \pi_{\Pi, k, n}(\mathbf{a}(x)) = \Delta_k^{\text{shift}} \right\}.$$

Then the set $\mathbb{E}_{\Pi}^{\text{cf}}$ is comeagre in $[0, 1]$ and $\dim_{\mathbb{P}}(\mathbb{E}_{\Pi}^{\text{cf}}) = 1$.

Proof This follows immediately by applying Theorem 2.3 and Theorem 2.4 to the IIFS in (5.2) noticing that in this case $\pi^{-1}(x) = \mathbf{a}(x)$ for all $x \in \mathbb{P}$. \square

Theorem 5.1 is a substantial extension of previously obtained results. For example, Šalat [30] proved that for a typical $x \in \mathbb{P}$, the oscillation of the frequencies $\pi_n(\mathbf{a}(x); i)$ is maximal for all digits i , i.e. for a typical $x \in \mathbb{P}$, we have

$$\liminf_n \pi_n(\mathbf{a}(x); i) = 0 \text{ and } \limsup_n \pi_n(\mathbf{a}(x); i) = 1 \text{ for all } i. \quad (5.4)$$

However, this result follows immediately from Theorem 5.1; indeed, if $\Pi = (\Pi_n)_n$ denotes the trivial averaging system with $\Pi_n = (\delta_{n,i})_i$ and \mathbb{M}^{cf} denotes the set of points x satisfying (5.4), then clearly $\mathbb{M}^{\text{cf}} \subseteq \mathbb{E}_{\Pi}^{\text{cf}}$, and since Theorem 5.1 shows that $\mathbb{E}_{\Pi}^{\text{cf}}$ is comeagre, we conclude that \mathbb{M}^{cf} is comeagre. Also, Madritsch [18] proved that for a typical $x \in \mathbb{P}$, the set of accumulation points of the m 'th order Hölder-average

of the frequencies of all strings of continued fraction digits of length k is maximal, i.e. for a typical $x \in \mathbb{P}$, we have

$$\operatorname{acc}_n \pi_{H,k,n}^m(\mathbf{a}(x)) = \Delta_k^{\text{shift}} \text{ for all } k \text{ and all } m$$

(recall, that for an infinite string $\mathbf{i} \in \mathbb{N}^{\mathbb{N}}$, the vector $\pi_{H,k,n}^m(\mathbf{i}) = (H_n^m(\pi(\mathbf{i}; \mathbf{l})))_{\mathbf{l} \in \mathbb{N}^k}$ of m 'th order Hölder-average frequencies $H_n^m(\pi(\mathbf{i}; \mathbf{l}))$ of all strings $\mathbf{l} \in I^k$ of length k is defined in (3.3)). However, this result follows immediately by applying Theorem 5.1 to the averaging systems Π_H^m (with $m \in \mathbb{N}$) in Sect. 3. Due to the importance and prevalence of the Hölder averages, we have decided to state this result formally in Corollary 5.2 below.

Corollary 5.2 *Write*

$$\mathbb{E}_H^{\text{cf}} = \bigcap_{m,k} \left\{ x \in \mathbb{P} \mid \operatorname{acc}_n \pi_{H,k,n}^m(\mathbf{a}(x)) = \Delta_k^{\text{shift}} \right\},$$

i.e. \mathbb{E}_H^{cf} is the set of points for which the set of accumulation points of the m 'th order Hölder-average of the frequencies of all strings of continued fraction digits of length k is maximal for all k and all m . Then the set \mathbb{E}_H^{cf} is comeagre in $[0, 1]$ and $\dim_{\mathbb{P}}(\mathbb{E}_H^{\text{cf}}) = 1$.

5.2 Lüroth expansion

As in the previous example, let \mathbb{P} denote the irrational numbers in the closed unit interval, i.e. $\mathbb{P} = [0, 1] \setminus \mathbb{Q}$. For each $x \in \mathbb{P}$, it is known that there are unique positive integers $b_1(x), b_2(x), \dots \in \mathbb{N} \setminus \{1\}$, such that

$$\begin{aligned} x = & \frac{1}{b_1(x)} + \frac{1}{b_1(x)(b_1(x)-1)b_2(x)} + \dots \\ & + \frac{1}{b_1(x)(b_1(x)-1) \cdots b_{n-1}(x)(b_{n-1}(x)-1)b_n(x)} + \dots \end{aligned} \quad (5.5)$$

The expansion in (5.5) is called the Lüroth expansion and the digits $b_1(x), b_2(x), \dots \in \mathbb{N} \setminus \{1\}$ are called the Lüroth digits of x ; see, for example, [6] for a more detailed discussion of the Lüroth expansion. We will also write $\mathbf{b}(x)$ for the string of the Lüroth digits of x , i.e.

$$\mathbf{b}(x) = b_1(x)b_2(x)\dots \in (\mathbb{N} \setminus \{1\})^{\mathbb{N}}.$$

In order to analyse the frequencies of the Lüroth digits of x , we define the IIFS $(X, (S_i)_{i \in I})$ as follows. Put $X = [0, 1]$, $I = \mathbb{N} \setminus \{1\}$ and for $i \in \mathbb{N} \setminus \{1\}$, define

$S_i : [0, 1] \rightarrow [0, 1]$ by

$$S_i(x) = \frac{x}{i(i-1)} + \frac{1}{i}. \quad (5.6)$$

It is well-known that $\pi((\mathbb{N} \setminus \{1\})^{\mathbb{N}}) = \mathbb{P}$ and that π is a homeomorphism between $(\mathbb{N} \setminus \{1\})^{\mathbb{N}}$ and \mathbb{P} . It is also well-known that $\pi^{-1}(x) = \mathbf{b}(x)$ for all $x \in \mathbb{P}$; it is, of course, this identity that allow us to apply Theorem 2.3 to the IIFS in (5.4) in order to obtain results about the Lüroth expansion digits of a typical number $x \in \mathbb{P}$.

As in the previous example, recall that if n is a positive integer and $i \in \mathbb{N}$ is a digit, then $\pi_n(\mathbf{i}; i) = \frac{|\{1 \leq j \leq n \mid i_j = i\}|}{n}$ denotes the frequency of the digit i among the first n digits of the string $\mathbf{i} = i_1 i_2 \dots$. In particular, $\pi_n(\mathbf{b}(x); i)$ denotes the frequency of the digit i among the first n digits of the Lüroth expansion digits of x . Since the Lebesgue measure is T -invariant and ergodic where $T : [0, 1] \rightarrow [0, 1]$ is defined by $T(x) = S_i^{-1}(x)$ for $x \in (\frac{1}{i}, \frac{1}{i-1}]$ and $T(0) = 0$, it follows from the ergodic theorem that for Lebesgue almost all $x \in \mathbb{P}$ we have

$$\pi_n(\mathbf{b}(x); i) \rightarrow \frac{1}{i-1} - \frac{1}{i} = \frac{1}{i(i-1)} \quad (5.7)$$

for all $i \in \mathbb{N}$. In analogy with normal numbers, we will say that a number $x \in \mathbb{P}$ is Lüroth normal if it satisfies (5.7). Hence, using this terminology, (5.7) says that Lebesgue almost all $x \in \mathbb{P}$ are Lüroth normal. As an application of Theorem 2.3 to the IIFS in (5.6), we will now show that the behaviour of the Lüroth expansion digits of a typical point x in \mathbb{P} is so irregular that *no* averaging system (regardless of how powerful it is) is powerful enough to “smoothen out” the behaviour of the frequencies of the digits; more precisely, for *all* averaging systems Π (regardless of how powerful they are), the set of accumulation points of the sequence $(\pi_{\Pi, k, n}(\mathbf{b}(x)))_n$ of Π -average frequencies of a typical number x is as big as possible. This is the content of the next theorem.

Theorem 5.3 *Let Π be an averaging system and write*

$$\mathbb{E}_{\Pi}^{\perp} = \bigcap_k \left\{ x \in \mathbb{P} \mid \operatorname{acc}_n \pi_{\Pi, k, n}(\mathbf{b}(x)) = \Delta_k^{\text{shift}} \right\}.$$

Then the set \mathbb{E}_{Π}^{\perp} is comeagre in $[0, 1]$ and $\dim_{\mathbb{P}}(\mathbb{E}_{\Pi}^{\perp}) = 1$.

Proof The proof of this result is similar to the proof of Theorem 5.1. \square

Theorem 5.3 is a substantial extension of previously obtained results. For example, Madritsch [18] proved that for a typical $x \in \mathbb{P}$, the set of accumulation points of the m 'th order Hölder-average of the frequencies of all strings of the Lüroth expansion digits of length k is maximal, i.e. for a typical $x \in \mathbb{P}$, we have

$$\operatorname{acc}_n \pi_{H, k, n}^m(\mathbf{b}(x)) = \Delta_k^{\text{shift}} \text{ for all } k \text{ and all } m.$$

However, this result follows immediately by applying Theorem 5.3 to the averaging systems $\Pi_{\mathbf{H}}^m$ (with $m \in \mathbb{N}$) in Sect. 3.

6 Proof of Theorem 2.3: Part 1

The purpose of this and the next section is to prove Theorem 2.3. The structure of the proof of Theorem 2.3 is as follows. For an averaging system Π , we first define a subset E_k of $I^{\mathbb{N}}$ such that

$$\mathbb{E}_{\Pi,k} = \pi(E_k)$$

(this is done in (6.1) below). The proof of Theorem 2.3 is now divided into the following two parts:

Part 1: In Part 1 we prove that the set E_k is comeagre in $I^{\mathbb{N}}$; this is done in this section (Sect. 6) and is the content of Proposition 6.6.

Part 2: In Part 2 we use the comeagreness of E_k in $I^{\mathbb{N}}$ to prove that $\mathbb{E}_{\Pi,k} = \pi(E_k)$ is comeagre in $K = \pi(I^{\mathbb{N}})$; this is done in the next section (Sect. 7).

We first recall and introduce various notation which will be used in this and the remaining sections of the paper. For $\mathbf{i} = i_1 i_2 \dots \in I^{\mathbb{N}}$ and a positive integer n , recall that we write $\mathbf{i}|n = i_1 \dots i_n$ for the truncation of \mathbf{i} at the n 'th place. Also, for $\mathbf{l} = l_1 \dots l_n \in I^*$ and $\mathbf{i} = i_1 \dots i_m \in I^*$, we write $\mathbf{li} = l_1 \dots l_n i_1 \dots i_m$ for the concatenation of \mathbf{l} and \mathbf{i} . Similarly, for $\mathbf{l} = l_1 \dots l_n \in I^*$ and $\mathbf{i} = i_1 i_2 \dots \in I^{\mathbb{N}}$, we write $\mathbf{li} = l_1 \dots l_n i_1 i_2 \dots$ for the concatenation of \mathbf{l} and \mathbf{i} . Also, if $\mathbf{i} \in I^n$ is a finite string with length equal to n , we write $|\mathbf{i}| = n$ for the length of \mathbf{i} . Finally, for $\mathbf{l} = l_1 \dots l_n \in I^*$, we write $[\mathbf{l}]$ for the cylinder generated by \mathbf{l} , i.e.

$$[\mathbf{l}] = \left\{ \mathbf{i} \in I^{\mathbb{N}} \mid \mathbf{i}|n = \mathbf{l} \right\}.$$

Let $\Pi = (\Pi_n)_n$ be an averaging system. For a positive integer k , define the set E_k by

$$E_k = \left\{ \mathbf{i} \in I^{\mathbb{N}} \mid \operatorname{acc}_n \pi_{\Pi,k,n}(\mathbf{i}) = \Delta_k^{\text{shift}} \right\}, \quad (6.1)$$

and note that

$$\mathbb{E}_{\Pi,k} = \pi(E_k);$$

i.e. the set E_k is the “symbolic” counterpart of the set $\mathbb{E}_{\Pi,k}$.

We now turn towards the proof of the main result in this section, namely, Proposition 6.6 saying that E_k is comeagre in $I^{\mathbb{N}}$. We first note that since $I^{\mathbb{N}}$ is a Baire space it suffices to find a countable family of open and dense subsets of $I^{\mathbb{N}}$ whose intersection is contained in E_k . Below we construct a family of subsets of $I^{\mathbb{N}}$ with these properties. Let $\Delta_k^{\text{shift,fin}}$ denote the family of shift invariant probability vectors in Δ_k^{shift} with only

finitely many non-zero components, i.e.

$$\Delta_k^{\text{shift,fin}} = \left\{ \mathbf{p} = (p_l)_{l \in \mathbb{N}^k} \in \Delta_k^{\text{shift}} \mid \text{only finitely many of the } p_l \text{'s are non-zero} \right\}.$$

It is clear that $\Delta_k^{\text{shift,fin}}$ is a dense and separable subset of Δ_k^{shift} , and we can therefore choose a countable dense subset Q_k of $\Delta_k^{\text{shift,fin}}$. For $\mathbf{q} \in Q_k$, $\varepsilon \in \mathbb{Q}^+$ and $m \in \mathbb{N}$, let

$$G_{\mathbf{q},\varepsilon,m} = \left\{ \mathbf{i} \in I^{\mathbb{N}} \mid \|\pi_{\Pi,k,n}(\mathbf{i}) - \mathbf{q}\|_1 < \varepsilon \text{ for some } n \geq m \right\}.$$

The family $(G_{\mathbf{q},\varepsilon,m})_{\mathbf{q} \in Q_k, \varepsilon \in \mathbb{Q}^+, m \in \mathbb{N}}$ is clearly a countable collection of subsets of $I^{\mathbb{N}}$. In order to show that E_k is comeagre it therefore suffices to show that: (1) the intersection of the sets $G_{\mathbf{q},\varepsilon,m}$ is contained in E_k , (2) the sets $G_{\mathbf{q},\varepsilon,m}$ are open in $I^{\mathbb{N}}$, and (3) the sets $G_{\mathbf{q},\varepsilon,m}$ are dense in $I^{\mathbb{N}}$; these statements are the contents of Proposition 6.1, Proposition 6.2 and Proposition 6.5 below.

Proposition 6.1 *We have $\bigcap_{\mathbf{q} \in Q_k} \bigcap_{\varepsilon \in \mathbb{Q}^+} \bigcap_{m \in \mathbb{N}} G_{\mathbf{q},\varepsilon,m} \subseteq E_k$.*

Proof Let $\overline{Q_k}$ denote the closure of Q_k in Δ_k with respect to the 1-norm. We clearly have

$$\bigcap_{\mathbf{q} \in Q_k} \bigcap_{\varepsilon \in \mathbb{Q}^+} \bigcap_{m \in \mathbb{N}} G_{\mathbf{q},\varepsilon,m} = \left\{ \mathbf{i} \in I^{\mathbb{N}} \mid \text{acc}_n \pi_{\Pi,k,n}(\mathbf{i}) = \overline{Q_k} \right\}. \quad (6.2)$$

However, since Q_k is dense in Δ_k^{shift} , we conclude that $\Delta_k^{\text{shift}} \subseteq \overline{Q_k}$, and it therefore follows from (6.2) and the fact that $\text{acc}_n \pi_{\Pi,k,n}(\mathbf{i}) \subseteq \Delta_k^{\text{shift}}$ for all $\mathbf{i} \in I^{\mathbb{N}}$ (by Theorem 2.1), that

$$\begin{aligned} \bigcap_{\mathbf{q} \in Q_k} \bigcap_{\varepsilon \in \mathbb{Q}^+} \bigcap_{m \in \mathbb{N}} G_{\mathbf{q},\varepsilon,m} &\subseteq \left\{ \mathbf{i} \in I^{\mathbb{N}} \mid \text{acc}_n \pi_{\Pi,k,n}(\mathbf{i}) = \Delta_k^{\text{shift}} \right\} \\ &= E_k. \end{aligned}$$

This completes the proof of Proposition 6.1. \square

Proposition 6.2 *The set $G_{\mathbf{q},\varepsilon,m}$ is open in $I^{\mathbb{N}}$.*

Proof Let $\mathbf{i} \in G_{\mathbf{q},\varepsilon,m}$. We must now show that there is $\mathbf{u} \in I^*$ such that $\mathbf{i} \in [\mathbf{u}]$ and $[\mathbf{u}] \subseteq G_{\mathbf{q},\varepsilon,m}$. Since $\mathbf{i} \in G_{\mathbf{q},\varepsilon,m}$, there is an integer n with $n \geq m$ such that $\|\pi_{\Pi,k,n}(\mathbf{i}) - \mathbf{q}\|_1 < \varepsilon$. Write $s_n = \max\{i \in \mathbb{N} \mid \pi_{n,i} \neq 0\}$ and note that it follows from the definition of an averaging system that $s_n < \infty$. Let $\mathbf{u} = \mathbf{i}|(s_n + k - 1) \in I^*$. We will prove that $\mathbf{i} \in [\mathbf{u}]$ and $[\mathbf{u}] \subseteq G_{\mathbf{q},\varepsilon,m}$. Indeed, it is clear that $\mathbf{i} \in [\mathbf{u}]$. Next, we prove that

$$[\mathbf{u}] \subseteq G_{\mathbf{q},\varepsilon,m}. \quad (6.3)$$

To prove (6.3), let $\mathbf{j} \in [\mathbf{u}]$. Since $\mathbf{i}, \mathbf{j} \in [\mathbf{u}]$, we conclude that $\pi_i(\mathbf{i}; \mathbf{l}) = \pi_i(\mathbf{j}; \mathbf{l})$ for all integers i with $1 \leq i \leq s_n$ and all $\mathbf{l} \in \mathbb{N}^k$. This clearly implies that

$$\begin{aligned} \boldsymbol{\pi}_{\Pi, k, n}(\mathbf{j}) &= (\pi_{\Pi, n}(\mathbf{j}; \mathbf{l}))_{\mathbf{l} \in \mathbb{N}^k} \\ &= \left(\sum_{i=1}^{s_n} \pi_{n, i} \pi_i(\mathbf{j}; \mathbf{l}) \right)_{\mathbf{l} \in \mathbb{N}^k} \\ &= \left(\sum_{i=1}^{s_n} \pi_{n, i} \pi_i(\mathbf{i}; \mathbf{l}) \right)_{\mathbf{l} \in \mathbb{N}^k} \\ &= (\pi_{\Pi, n}(\mathbf{i}; \mathbf{l}))_{\mathbf{l} \in \mathbb{N}^k} \\ &= \boldsymbol{\pi}_{\Pi, k, n}(\mathbf{i}), \end{aligned}$$

whence $\|\boldsymbol{\pi}_{\Pi, k, n}(\mathbf{j}) - \mathbf{q}\|_1 = \|\boldsymbol{\pi}_{\Pi, k, n}(\mathbf{i}) - \mathbf{q}\|_1 < \varepsilon$, and so $\mathbf{j} \in G_{\mathbf{q}, \varepsilon, m}$ because $n \geq m$. \square

Finally, we must prove that $G_{\mathbf{q}, \varepsilon, m}$ is dense in $I^{\mathbb{N}}$. In order to prove this we first prove the following auxiliary result, namely, if $\mathbf{q} \in \Delta_k^{\text{shift}, \text{fin}}$, then there is a finite subset $I_{\mathbf{q}}$ of I and $\mathbf{i} \in I_{\mathbf{q}}^{\mathbb{N}}$ such that $\|\boldsymbol{\pi}_{\Pi, k, n}(\mathbf{i}) - \mathbf{q}\|_1 \rightarrow 0$; this is the statement of Proposition 6.4. However, in order to prove Proposition 6.4 we need Theorem 6.3 below. Fix a positive integer N . For $x \in [0, 1]$, let

$$x = \sum_{n=1}^{\infty} \frac{d_{N, n}(x)}{N^n},$$

where $d_{N, n}(x) \in \{0, 1, \dots, N-1\}$ for $n \in \mathbb{N}$, denote the unique non-terminating N -adic expansion of x . For a positive integer n and a finite string $\mathbf{l} = l_1 \dots l_k \in \{0, 1, \dots, N-1\}^k$, we write

$$\Lambda_{N, k, n}(x; \mathbf{l}) = \frac{|\{1 \leq i \leq n \mid d_{N, i}(x) = l_1, \dots, d_{N, i+k-1}(x) = l_k\}|}{n}$$

for the frequency of the string \mathbf{l} among the first n digits in the N -adic expansion of x , and let

$$\mathbf{\Lambda}_{N, k, n}(x) = (\Lambda_{N, k, n}(x; \mathbf{l}))_{\mathbf{l} \in \{0, 1, \dots, N-1\}^k}$$

denote the vector of frequencies $\Lambda_{N, k, n}(x; \mathbf{l})$ of all strings $\mathbf{l} \in \{0, 1, \dots, N-1\}^k$ of length k . Let

$$\Delta_{N, k} = \left\{ (p_{\mathbf{l}})_{\mathbf{l} \in \{0, 1, \dots, N-1\}^k} \mid p_{\mathbf{l}} \geq 0, \sum_{\mathbf{l}} p_{\mathbf{l}} = 1 \right\},$$

and

$$\Delta_{N,k}^{\text{shift}} = \left\{ (p_{\mathbf{l}})_{\mathbf{l} \in \{0,1,\dots,N-1\}^k} \mid p_{\mathbf{l}} \geq 0, \sum_{\mathbf{l}} p_{\mathbf{l}} = 1, \right. \\ \left. \sum_{\mathbf{l}} p_{\mathbf{l}} = \sum_{\mathbf{l}'} p_{\mathbf{l}'} \text{ for all } \mathbf{l} \in \{0,1,\dots,N-1\}^{k-1} \right\},$$

i.e. $\Delta_{N,k}^{\text{shift}}$ ($\Delta_{N,k}^{\text{shift}}$) denotes the simplex of (shift invariant) probability vectors indexed by strings $\mathbf{l} = l_1 \dots l_k$ of length k with entries $l_j \in \{0, 1, \dots, N-1\}$. Finally, define $H_{N,k} : \Delta_{N,k} \rightarrow \mathbb{R}$ by

$$H_{N,k}(\mathbf{p}) = -\frac{1}{\log N} \sum_{\mathbf{l} \in \{0,1,\dots,N-1\}^{k-1}} \sum_{\mathbf{l}'} p_{\mathbf{l}'} \log \frac{p_{\mathbf{l}'}}{\sum_i p_{\mathbf{l}i}}.$$

for $\mathbf{p} = (p_{\mathbf{l}})_{\mathbf{l} \in \{0,1,\dots,N-1\}^k}$ (as usual, we put $0 \log 0 = 0$). We can now state Theorem 6.3; we note that this result is proved in [24, Theorem 1]. Recall, that \dim_{H} denotes the Hausdorff dimension.

Theorem 6.3 [24, Theorem 1] *Let $\mathbf{p} \in \Delta_{N,k}$.*

(1) *If $\mathbf{p} \notin \Delta_{N,k}^{\text{shift}}$, then*

$$\left\{ x \in [0, 1] \mid \lim_n \|\mathbf{A}_{N,k,n}(x) - \mathbf{p}\|_1 = 0 \right\} = \emptyset.$$

(2) *If $\mathbf{p} \in \Delta_{N,k}^{\text{shift}}$, then*

$$\dim_{\text{H}} \left\{ x \in [0, 1] \mid \lim_n \|\mathbf{A}_{N,k,n}(x) - \mathbf{p}\|_1 = 0 \right\} = H_{N,k}(\mathbf{p}).$$

Next, we can state and prove Proposition 6.4.

Proposition 6.4 *Let $\mathbf{q} \in \Delta_k^{\text{shift}, \text{fin}}$. Then there is a finite subset $I_{\mathbf{q}}$ of I and $\mathbf{i} \in I_{\mathbf{q}}^{\mathbb{N}}$ such that*

$$\|\pi_{\Pi,k,n}(\mathbf{i}) - \mathbf{q}\|_1 \rightarrow 0.$$

Proof Write $\mathbf{q} = (q_{\mathbf{l}})_{\mathbf{l} \in I^k}$. Since there are only finitely many $\mathbf{l} \in I^k$ such that $q_{\mathbf{l}} > 0$, we can find a finite subset $I_{\mathbf{q}} = \{i_{\mathbf{q},1}, \dots, i_{\mathbf{q},N}\}$ of I such that $q_{\mathbf{l}} = 0$ for $\mathbf{l} \in I^k \setminus I_{\mathbf{q}}^k$. For $\mathbf{l} = l_1 \dots l_k \in \{0, 1, \dots, N-1\}^k$, write $\tilde{\mathbf{l}} = \tilde{l}_1 \dots \tilde{l}_k \in I_{\mathbf{q}}^k$ where $\tilde{l}_j = i_{\mathbf{q},l_j+1}$, and let $\tilde{\mathbf{q}} = (q_{\tilde{\mathbf{l}}})_{\tilde{\mathbf{l}} \in \{0,1,\dots,N-1\}^k}$. Since $\mathbf{q} = (q_{\mathbf{l}})_{\mathbf{l} \in I^k} \in \Delta_k^{\text{shift}}$ and $q_{\mathbf{l}} = 0$ for $\mathbf{l} \in I^k \setminus I_{\mathbf{q}}^k$, we conclude that $\tilde{\mathbf{q}} \in \Delta_{N,k}^{\text{shift}}$, and it therefore follows immediately from Theorem 6.3 that $\dim_{\text{H}} \{x \in [0, 1] \mid \lim_n \|\mathbf{A}_{N,k,n}(x) - \tilde{\mathbf{q}}\|_1 = 0\} = H_{N,k}(\tilde{\mathbf{q}}) > 0$, and so $\{x \in [0, 1] \mid \lim_n \|\mathbf{A}_{N,k,n}(x) - \tilde{\mathbf{q}}\|_1 = 0\} \neq \emptyset$. In particular, it follows that we can thus

choose $x \in [0, 1]$ such that $\lim_n \|\mathbf{\Lambda}_{N,k,n}(x) - \tilde{\mathbf{q}}\|_1 = 0$. Next, put $\mathbf{i} = i_1 i_2 \dots \in I^{\mathbb{N}}$ where $i_j = i_{\mathbf{q}, d_{N,j}(x)+1}$. We now claim that

$$\|\pi_n(\mathbf{i}; k) - \mathbf{q}\|_1 \rightarrow 0. \quad (6.4)$$

Indeed, since $\lim_n \|\mathbf{\Lambda}_{N,k,n}(x) - \tilde{\mathbf{q}}\|_1 = 0$, we deduce that $\Lambda_{N,k,n}(x; \mathbf{l}) \rightarrow q_{\tilde{\mathbf{l}}}$ for all $\mathbf{l} \in \{0, 1, \dots, N-1\}^k$. It follows immediately from this that $\pi_n(\mathbf{i}; \mathbf{l}) \rightarrow q_{\mathbf{l}}$ for all $\mathbf{l} \in I_{\tilde{\mathbf{q}}}^k$, and the Consistency Condition therefore implies that $\sum_i \pi_{n,i} \pi_i(\mathbf{i}; \mathbf{l}) \rightarrow q_{\mathbf{l}}$ for all $\mathbf{l} \in I_{\tilde{\mathbf{q}}}^k$, whence

$$\forall \mathbf{l} \in I_{\tilde{\mathbf{q}}}^k : \left| \sum_i \pi_{n,i} \pi_i(\mathbf{i}; \mathbf{l}) - q_{\mathbf{l}} \right| \rightarrow 0. \quad (6.5)$$

Since $I_{\mathbf{q}}$ is finite, we deduce from (6.5) that

$$\sum_{\mathbf{l} \in I_{\tilde{\mathbf{q}}}^k} \left| \sum_i \pi_{n,i} \pi_i(\mathbf{i}; \mathbf{l}) - q_{\mathbf{l}} \right| \rightarrow 0. \quad (6.6)$$

Also, since $\mathbf{i} \in I_{\tilde{\mathbf{q}}}^{\mathbb{N}}$ and $q_{\mathbf{l}} = 0$ for all $\mathbf{l} \in I^k \setminus I_{\tilde{\mathbf{q}}}^k$, we conclude that $\pi_n(\mathbf{i}; \mathbf{l}) = q_{\mathbf{l}} = 0$ for all $\mathbf{l} \in I^k \setminus I_{\tilde{\mathbf{q}}}^k$ and all n , and so

$$\forall \mathbf{l} \in I^k \setminus I_{\tilde{\mathbf{q}}}^k : \left| \sum_i \pi_{n,i} \pi_i(\mathbf{i}; \mathbf{l}) - q_{\mathbf{l}} \right| = 0 \text{ for all } n,$$

whence

$$\sum_{\mathbf{l} \in I^k \setminus I_{\tilde{\mathbf{q}}}^k} \left| \sum_i \pi_{n,i} \pi_i(\mathbf{i}; \mathbf{l}) - q_{\mathbf{l}} \right| = 0 \text{ for all } n, \quad (6.7)$$

Finally, it follows from (6.6) and (6.7) that $\|\pi_{\Pi,k,n}(\mathbf{i}) - \mathbf{q}\|_1 = \sum_{\mathbf{l} \in I^k} \left| \sum_i \pi_{n,i} \pi_i(\mathbf{i}; \mathbf{l}) - q_{\mathbf{l}} \right| \rightarrow 0$. This completes the proof of (6.4). \square

Proposition 6.5 *The set $G_{\mathbf{q},\varepsilon,m}$ is dense in $I^{\mathbb{N}}$.*

Proof It follows from Proposition 6.4 that there is a finite subset $I_{\mathbf{q}}$ of I and $\mathbf{i} \in I_{\tilde{\mathbf{q}}}^{\mathbb{N}}$ such that $\|\pi_{\Pi,k,n}(\mathbf{i}) - \mathbf{q}\|_1 \rightarrow 0$. Now put $U = \cup_{\mathbf{u} \in I^*} [\mathbf{u}\mathbf{i}]$. We claim that U is dense in $I^{\mathbb{N}}$ and that $U \subseteq G_{\mathbf{q},\varepsilon,m}$. Indeed, it is clear that U is dense in $I^{\mathbb{N}}$. Next, we prove that $U \subseteq G_{\mathbf{q},\varepsilon,m}$. In order to prove this, we let $\mathbf{u} \in I^*$. We must now prove that $\mathbf{u}\mathbf{i} \in G_{\mathbf{q},\varepsilon,m}$. To prove this, we first prove the following two claims.

Claim 1 *We have $\|\pi_{\Pi,k,n}(\mathbf{u}\mathbf{i}) - \pi_{\Pi,k,n}(\mathbf{i})\|_1 \rightarrow 0$.*

Proof of Claim 1. First note that since $\mathbf{u} \in I^*$, we may choose a finite subset M of I such that $I_{\mathbf{q}} \subseteq M$ and $\mathbf{u} \in M^*$. Next, note that

$$\begin{aligned} \|\pi_{\Pi,k,n}(\mathbf{i}) - \pi_{\Pi,k,n}(\mathbf{ui})\|_1 &= \left| \sum_{\mathbf{l} \in I^k} \sum_i \pi_{n,i} \pi_i(\mathbf{i}; \mathbf{l}) - \sum_{\mathbf{l} \in I^k} \sum_i \pi_{n,i} \pi_i(\mathbf{ui}; \mathbf{l}) \right| \\ &\leq \sum_{\mathbf{l} \in I^k} \sum_i \pi_{n,i} |\pi_i(\mathbf{i}; \mathbf{l}) - \pi_i(\mathbf{ui}; \mathbf{l})| \\ &\leq \sum_{\mathbf{l} \in M^k} \sum_i \pi_{n,i} |\pi_i(\mathbf{i}; \mathbf{l}) - \pi_i(\mathbf{ui}; \mathbf{l})| \\ &\quad + \sum_{\mathbf{l} \in I^k \setminus M^k} \sum_i \pi_{n,i} |\pi_i(\mathbf{i}; \mathbf{l}) - \pi_i(\mathbf{ui}; \mathbf{l})|. \end{aligned} \quad (6.8)$$

We first analyse the sum $\sum_{\mathbf{i} \in M^k} \sum_i \pi_{n,i} |\pi_i(\mathbf{i}; \mathbf{l}) - \pi_i(\mathbf{ui}; \mathbf{l})|$. Indeed, we immediately observe that it is clear that for all positive integers n , we have

$$\begin{aligned} \forall \mathbf{l} \in M^k : \quad &|n\pi_n(\mathbf{i}; \mathbf{l}) - (n + |\mathbf{u}|)\pi_{n+|\mathbf{u}|}(\mathbf{ui}; \mathbf{l})| \leq |\mathbf{u}|, \\ \forall \mathbf{l} \in M^k : \quad &|n\pi_n(\mathbf{ui}; \mathbf{l}) - (n + |\mathbf{u}|)\pi_{n+|\mathbf{u}|}(\mathbf{ui}; \mathbf{l})| \leq |\mathbf{u}|. \end{aligned}$$

In particular, this implies that for all positive integers n , we have

$$\begin{aligned} \forall \mathbf{l} \in M^k : \quad &|\pi_n(\mathbf{i}; \mathbf{l}) - \pi_n(\mathbf{ui}; \mathbf{l})| \leq |\pi_n(\mathbf{i}; \mathbf{l}) - \pi_{n+|\mathbf{u}|}(\mathbf{ui}; \mathbf{l})| \\ &+ |\pi_{n+|\mathbf{u}|}(\mathbf{ui}; \mathbf{l}) - \pi_n(\mathbf{ui}; \mathbf{l})| \\ &\leq (1 + \pi_{n+|\mathbf{u}|}(\mathbf{ui}; \mathbf{l})) \frac{|\mathbf{u}|}{n} + (1 + \pi_{n+|\mathbf{u}|}(\mathbf{ui}; \mathbf{l})) \frac{|\mathbf{u}|}{n} \\ &\leq 4 \frac{|\mathbf{u}|}{n} \end{aligned}$$

(since $\pi_{n+|\mathbf{u}|}(\mathbf{ui}; \mathbf{l}) \leq 1$), and so $|\pi_n(\mathbf{i}; \mathbf{l}) - \pi_n(\mathbf{ui}; \mathbf{l})| \rightarrow 0$. It follows from this and the Consistency Condition that

$$\forall \mathbf{l} \in M^k : \quad \sum_i \pi_{n,i} |\pi_i(\mathbf{i}; \mathbf{l}) - \pi_i(\mathbf{ui}; \mathbf{l})| \rightarrow 0.$$

Finally, since the set M^k is finite, we therefore conclude that

$$\sum_{\mathbf{l} \in M^k} \sum_i \pi_{n,i} |\pi_i(\mathbf{i}; \mathbf{l}) - \pi_i(\mathbf{ui}; \mathbf{l})| \rightarrow 0. \quad (6.9)$$

Next, we analyse the sum $\sum_{\mathbf{i} \in I^k \setminus M^k} \sum_i \pi_{n,i} |\pi_i(\mathbf{i}; \mathbf{l}) - \pi_i(\mathbf{ui}; \mathbf{l})|$. Since $\mathbf{u} \in M^*$ and $\mathbf{i} \in I_{\mathbf{q}}^{\mathbb{N}} \subseteq M^{\mathbb{N}}$, the strings \mathbf{u} and \mathbf{i} only contain letters from M , and it therefore follows that if $\mathbf{l} \in I^k \setminus M^k$, then \mathbf{l} cannot be a substring of \mathbf{u} or \mathbf{i} . In particular, we

conclude that for all positive integers n , we have $\pi_n(\mathbf{i}; \mathbf{l}) = \pi_n(\mathbf{ui}; \mathbf{l}) = 0$ for all $\mathbf{l} \in I^k \setminus M^k$, and so

$$\forall \mathbf{l} \in I^k \setminus M^k : \sum_i \pi_{n,i} |\pi_i(\mathbf{i}; \mathbf{l}) - \pi_i(\mathbf{ui}; \mathbf{l})| = 0 \text{ for all } n.$$

This clearly implies that

$$\sum_{\mathbf{i} \in I^k \setminus M^k} \sum_i \pi_{n,i} |\pi_i(\mathbf{i}; \mathbf{l}) - \pi_i(\mathbf{ui}; \mathbf{l})| = 0 \text{ for all } n. \quad (6.10)$$

The statement in Claim 1 now follows immediately from combining (6.8), (6.9) and (6.10). This completes the proof of Claim 1. \square

Claim 2 We have $\|\pi_{\Pi,k,n}(\mathbf{ui}) - \mathbf{q}\|_1 \rightarrow 0$.

Proof of Claim 2. Indeed, since $\|\pi_{\Pi,k,n}(\mathbf{i}) - \mathbf{q}\|_1 \rightarrow 0$, it follows immediately from Claim 1 that $\|\pi_{\Pi,k,n}(\mathbf{ui}) - \mathbf{q}\|_1 \leq \|\pi_{\Pi,k,n}(\mathbf{ui}) - \pi_{\Pi,k,n}(\mathbf{i})\|_1 + \|\pi_{\Pi,k,n}(\mathbf{i}) - \mathbf{q}\|_1 \rightarrow 0$. This completes the proof of Claim 2.

Finally, it follows immediately from Claim 2 that $\mathbf{ui} \in G_{\mathbf{q},\varepsilon,m}$. \square

We can now prove the main result in this section, namely, that the set E_k is comeagre in $I^{\mathbb{N}}$.

Proposition 6.6 *The set E_k is comeagre in $I^{\mathbb{N}}$*

Proof Indeed, it follows immediately from Proposition 6.1, Proposition 6.2 and Proposition 6.5 that the family $(G_{\mathbf{q},\varepsilon,m})_{\mathbf{q} \in Q_k, \varepsilon \in \mathbb{Q}^+, m \in \mathbb{N}}$ is a countable collection of open and dense sets whose union is contained in E_k . We conclude immediately from this that E_k is comeagre. \square

7 Proof of Theorem 2.3: Part 2

The main purpose of this section is to prove Theorem 2.3 using Proposition 6.6. We first prove the following three auxiliary results. The first lemma, i.e. Lemma 7.1, is well-known if the index set I is finite, see, for example [8]; however, we have been unable to find a proof in the general case, and we have therefore decided to include the simple and brief proof. Recall, that $K = \pi(I^{\mathbb{N}})$ denotes the self-similar set associated with the IIFS $(X, (S_i)_{i \in I})$.

Lemma 7.1 *Let $(X, (S_i)_{i \in I})$ be an IIFS. Assume that $(X, (S_i)_{i \in I})$ satisfies the SOSC with open set V . Then $K \subseteq \overline{V}$.*

Proof Let $\mathbf{i} \in I^{\mathbb{N}}$. We will now prove that $\pi(\mathbf{i}) \in \overline{V}$. The continuity of S_i implies that $S_i(\overline{V}) \subseteq \overline{S_i(V)}$, whence $S_i(\overline{V}) \subseteq \overline{S_i(V)} \subseteq \overline{V}$. It follows from this that $(S_{i|n}(\overline{V}))_n$ is a decreasing sequence of non-empty compact sets with $\text{diam}(S_{i|n}(\overline{V})) \rightarrow 0$, and the intersection $\cap_n S_{i|n}(\overline{V})$ is therefore a singleton. However, since $\overline{V} \subseteq X$, we conclude that $\cap_n S_{i|n}(\overline{V}) \subseteq \cap_n S_{i|n}(X) = \{\pi(\mathbf{i})\}$. It follows immediately from this that $\{\pi(\mathbf{i})\} = \cap_n S_{i|n}(\overline{V})$, and so $\pi(\mathbf{i}) \in \cap_n S_{i|n}(\overline{V}) \subseteq \overline{V}$. \square

Lemma 7.2 *Let $(X, (S_i)_{i \in I})$ be an IIFS. Assume that $(X, (S_i)_{i \in I})$ satisfies the SOSC with open set V and let $H \subseteq I^{\mathbb{N}}$. If $\mathbf{l}, \mathbf{k} \in I^*$ satisfy $[\mathbf{l}] \cap H = \emptyset$ and $S_{\mathbf{k}}(K) \subseteq V$, then $\pi([\mathbf{l}\mathbf{k}]) \cap \pi(H) = \emptyset$.*

Proof Assume, in order to reach a contradiction, that $\pi([\mathbf{l}\mathbf{k}]) \cap \pi(H) \neq \emptyset$. We can thus choose $x \in \pi([\mathbf{l}\mathbf{k}]) \cap \pi(H)$. Since $x \in \pi([\mathbf{l}\mathbf{k}])$, there is $\mathbf{r} \in I^{\mathbb{N}}$ such that

$$x = \pi(\mathbf{lkr}).$$

Also, since $x \in \pi(H)$, we can find $\mathbf{m} \in I^*$ with $|\mathbf{m}| = |\mathbf{l}|$ and $\mathbf{s} \in I^{\mathbb{N}}$ such that $\mathbf{ms} \in H$ and

$$x = \pi(\mathbf{ms}).$$

Observe that $\mathbf{m} \neq \mathbf{l}$; indeed, otherwise $\mathbf{m} = \mathbf{l}$, whence $\mathbf{ms} \in [\mathbf{l}] \cap H$, contradicting the fact that $[\mathbf{l}] \cap H = \emptyset$.

Since $|\mathbf{m}| = |\mathbf{l}|$ and $\mathbf{m} \neq \mathbf{l}$, we conclude that $S_{\mathbf{l}}(V) \cap S_{\mathbf{m}}(V) = \emptyset$. Next, recall that the maps S_i for $i \in I$ are assumed to be injective, and since the maps S_i are also continuous (because they are Lipschitz), we conclude from the Invariance of Domain that each map S_i is open. It follows from this and the fact that V is open, that the sets $S_{\mathbf{l}}(V)$ and $S_{\mathbf{m}}(V)$ are open. Hence, the sets $S_{\mathbf{l}}(V)$ and $S_{\mathbf{m}}(V)$ are open with $S_{\mathbf{l}}(V) \cap S_{\mathbf{m}}(V) = \emptyset$. It is not difficult to see that this implies that

$$S_{\mathbf{l}}(V) \cap \overline{S_{\mathbf{m}}(V)} = \emptyset. \quad (7.1)$$

However, we also clearly have (since, by assumption, $S_{\mathbf{k}}(K) \subseteq V$)

$$x = \pi(\mathbf{lkr}) \in S_{\mathbf{l}\mathbf{k}}(K) = S_{\mathbf{l}}S_{\mathbf{k}}(K) \subseteq S_{\mathbf{l}}(V), \quad (7.2)$$

and (since $K \subseteq \overline{V}$ by Lemma 7.1)

$$x = \pi(\mathbf{ms}) \in S_{\mathbf{m}}(K) \subseteq S_{\mathbf{m}}(\overline{V}) = \overline{S_{\mathbf{m}}(V)}. \quad (7.3)$$

The desired contradiction now follows from (7.1), (7.2) and (7.3). \square

Proposition 7.3 *Let $(X, (S_i)_{i \in I})$ be an IIFS. Assume that $(X, (S_i)_{i \in I})$ satisfies the SOSC and that the map π is closed. If $M \subseteq I^{\mathbb{N}}$ is meagre in $I^{\mathbb{N}}$, then $\pi(M)$ is meagre in $K = \pi(I^{\mathbb{N}})$.*

Proof Since $M \subseteq I^{\mathbb{N}}$ is meagre in $I^{\mathbb{N}}$, we can find countably many subsets N_1, N_2, \dots of $I^{\mathbb{N}}$ such that

$$M = \bigcup_n N_n \quad (7.4)$$

and such that each set N_n is nowhere dense in $I^{\mathbb{N}}$, i.e.

$$\overline{N_n}^{\circ} = \emptyset \quad (7.5)$$

for all positive integers n . We now prove the following two claims.

Claim 1. *Let n be a positive integer. If $\mathbf{i} \in I^*$ satisfies $\pi([\mathbf{i}]) \subseteq \pi(\overline{N_n})$, then $[\mathbf{i}] \subseteq \overline{N_n}$.*

Proof of Claim 1. Assume, in order to reach a contradiction, that $[\mathbf{i}] \not\subseteq \overline{N_n}$. We can thus choose $\mathbf{r} \in I^{\mathbb{N}}$ with $\mathbf{ir} \in [\mathbf{i}] \cap (I^{\mathbb{N}} \setminus \overline{N_n})$, and since $[\mathbf{i}] \cap (I^{\mathbb{N}} \setminus \overline{N_n})$ is open, this implies that there exists a positive integer n such that if $\mathbf{j} = \mathbf{r}|n$, then $[\mathbf{ij}] \subseteq [\mathbf{i}] \cap (I^{\mathbb{N}} \setminus \overline{N_n})$, i.e.

$$[\mathbf{ij}] \cap \overline{N_n} = \emptyset. \quad (7.6)$$

Next, since $(X, (S_i)_{i \in I})$ satisfies the SOSC there is a non-empty, open and bounded subset V of X with $K \cap V \neq \emptyset$ such that $S_i(V) \subseteq V$ for all i and $S_i(V) \cap S_j(V) = \emptyset$ for all i, j with $i \neq j$. Because $K \cap V \neq \emptyset$, there is $\mathbf{w} \in I^{\mathbb{N}}$ such that $\pi(\mathbf{w}) \in V$. In particular, this shows that $\cap_m S_{\mathbf{w}|m}(K) \subseteq \cap_m S_{\mathbf{w}|m}(X) = \{\pi(\mathbf{w})\} \subseteq V$, and since V is open and $\lim_m \text{diam}(S_{\mathbf{w}|m}(K)) = 0$, we therefore conclude that there is a positive integer m such that $S_{\mathbf{w}|m}(K) \subseteq V$. Putting $\mathbf{k} = \mathbf{w}|m \in I^*$, we therefore deduce that

$$S_{\mathbf{k}}(K) \subseteq V. \quad (7.7)$$

It now follows from (7.6), (7.7) and Lemma 7.2 that

$$\pi([\mathbf{ijk}]) \cap \pi(\overline{N_n}) = \emptyset.$$

However, this clearly contradicts the fact that $\pi([\mathbf{ijk}]) \subseteq \pi([\mathbf{i}]) \subseteq \pi(\overline{N_n})$. This completes the proof of Claim 1.

Claim 2. *Let n be a positive integer. Then $\pi(N_n)$ is nowhere dense in K , i.e. $\overline{\pi(N_n)}^\circ = \emptyset$.*

Proof of Claim 2. Assume, in order to reach a contradiction, that $\overline{\pi(N_n)}^\circ \neq \emptyset$. This implies that $\pi^{-1}(\overline{\pi(N_n)}^\circ) \neq \emptyset$. It follows from this and the fact that $\pi^{-1}(\overline{\pi(N_n)}^\circ)$ is open that we can choose a string $\mathbf{i} \in I^*$ such that $[\mathbf{i}] \subseteq \pi^{-1}(\overline{\pi(N_n)}^\circ)$, whence

$$\pi([\mathbf{i}]) \subseteq \overline{\pi(N_n)}^\circ \subseteq \overline{\pi(N_n)}. \quad (7.8)$$

However, since the map π is assumed to be closed, we deduce that $\pi(\overline{N_n})$ is closed, i.e. $\pi(\overline{N_n}) = \overline{\pi(N_n)}$. Since also $\pi(N_n) \subseteq \pi(\overline{N_n})$, we therefore conclude that

$$\overline{\pi(N_n)} \subseteq \overline{\pi(\overline{N_n})} = \pi(\overline{N_n}). \quad (7.9)$$

Combining (7.8) and (7.9), we see that $\pi([\mathbf{i}]) \subseteq \pi(\overline{N_n})$, and it therefore follows from Claim 1 that $[\mathbf{i}] \subseteq \overline{N_n}$, whence $[\mathbf{i}] = [\mathbf{i}]^\circ \subseteq \overline{N_n}^\circ$, and so

$$\overline{N_n}^\circ \neq \emptyset. \quad (7.10)$$

The desired contradiction now follows by comparing (7.5) and (7.10). This completes the proof of Claim 2.

We can now prove that $\pi(M)$ is comeagre in K . Indeed, it follows from (7.4) that

$$\pi(M) = \bigcup_n \pi(N_n)$$

and since Claim 2 shows that $\pi(N_n)$ is nowhere dense in K for all positive integers n , we conclude that $\pi(M)$ is meagre in K . \square

We can now prove Theorem 2.3

Proof of Theorem 2.3 It follows from Proposition 6.6 that the set $I^N \setminus E_k$ is meagre in I^N and Proposition 7.3 therefore shows that the set $\pi(I^N \setminus E_k)$ is meagre in K . However, since $K \setminus \pi(E_k) \subseteq \pi(I^N \setminus E_k)$, we conclude from this that $K \setminus \pi(E_k)$ is meagre in K , i.e. $\pi(E_k)$ is comeagre in K .

8 Proof of Theorem 2.4

The purpose of this section is to prove Theorem 2.4. We first prove several auxiliary results. Recall that \dim_P denotes the packing dimension.

Lemma 8.1 *Let $E, X \subseteq \mathbb{R}^d$ with $E \subseteq X$ and $\dim_P(E) < \dim_P(X)$. Assume that $\dim_P(X \cap U) = \dim_P(X)$ for all open subsets U of \mathbb{R}^d with $X \cap U \neq \emptyset$. Then E is meagre in X .*

Proof For $F \subseteq \mathbb{R}^d$, let $\overline{\dim}_B(F)$ denote the upper box-dimension of F ; the reader is referred to [8] for the definition of the upper box-dimension $\overline{\dim}_B$. Also, recall that if $F \subseteq \mathbb{R}^d$, then

$$\dim_P(F) = \inf \left\{ \sup_n \overline{\dim}_B(E_n) \mid E_1, E_2, \dots \subseteq \mathbb{R}^d \text{ and } F \subseteq \bigcup_n E_n \right\};$$

see, for example, [8]. Since $\dim_P(E) < \dim_P(X)$ it therefore follows that we can find countably many subsets E_1, E_2, \dots of \mathbb{R}^d with $E \subseteq \bigcup_n E_n$ such that $\sup_n \overline{\dim}_B(E_n) < \dim_P(X)$.

Since $E \subseteq X$ and $E \subseteq \bigcup_n E_n$, we immediately conclude that

$$E \subseteq \bigcup_n (X \cap E_n). \quad (8.1)$$

Next, we prove the following statement:

$$\text{the set } X \cap E_n \text{ is nowhere dense in } X \text{ for all } n. \quad (8.2)$$

We will now prove (8.2). Assume in order to reach a contradiction that the statement in (8.2) is not true. We can therefore find an integer n such that $X \cap E_n$ is not nowhere dense in X . This implies that there is an open subset U of \mathbb{R}^d such that $X \cap U \neq \emptyset$ and

$X \cap E_n \cap U$ is dense in $X \cap U$. In particular, we deduce that $X \cap U \subseteq \overline{X \cap E_n \cap U}$, whence

$$\begin{aligned} \overline{\dim}_B(E_n) &= \overline{\dim}_B(\overline{E_n}) \\ &\geq \overline{\dim}_B(\overline{X \cap E_n \cap U}) \\ &\geq \overline{\dim}_B(X \cap U) \\ &\geq \dim_P(X \cap U) \\ &= \dim_P(X); \end{aligned} \quad (8.3)$$

in the first line of (8.3) we have used the fact that if $F \subseteq \mathbb{R}^d$, then $\overline{\dim}_B(F) = \overline{\dim}_B(\overline{F})$, and in the fourth line of (8.3) we have used the fact that if $F \subseteq \mathbb{R}^d$, then $\overline{\dim}_B(F) \geq \dim_P(F)$. However, since $\sup_m \overline{\dim}_B(E_m) < \dim_P(X)$, we also conclude that

$$\overline{\dim}_B(E_n) \leq \sup_m \overline{\dim}_B(E_m) < \dim_P(X). \quad (8.4)$$

The desired contradiction now follows from (8.3) and (8.4). This completes the proof of (8.2).

Finally, it follows immediately from (8.1) and (8.2) that E is meagre in X . \square

Lemma 8.2 *Let $(X, (S_i)_{i \in I})$ be an IIFS. Assume that the maps S_i are bi-Lipschitz for all i . Then $\dim_P(K \cap U) = \dim_P(K)$ for all open subsets U of \mathbb{R}^d with $K \cap U \neq \emptyset$.*

Proof It is clear that $\dim_P(K \cap U) \leq \dim_P(K)$. Next, we prove the reverse inequality. Because $K \cap U \neq \emptyset$, we can choose a string $\mathbf{i} \in I^{\mathbb{N}}$ with $\pi(\mathbf{i}) \in U$. In particular, we have $\cap_n S_{\mathbf{i}|n}(K) \subseteq \cap_n S_{\mathbf{i}|n}(X) = \{\pi(\mathbf{i})\} \subseteq U$, and since U is open and $\lim_n \text{diam}(S_{\mathbf{i}|n}(K)) = 0$, it therefore follows that we can find a positive integer n with $S_{\mathbf{i}|n}(K) \subseteq U$, whence $S_{\mathbf{i}|n}(K) \subseteq K \cap U$. Also, since $S_{\mathbf{i}|n}$ is bi-Lipschitz, we conclude that $\dim_P(S_{\mathbf{i}|n}(K)) = \dim_P(K)$, and it follows from this and the inclusion $S_{\mathbf{i}|n}(K) \subseteq K \cap U$ that $\dim_P(K) = \dim_P(S_{\mathbf{i}|n}(K)) \leq \dim_P(K \cap U)$. \square

Lemma 8.3 *Let $(X, (S_i)_{i \in I})$ be an IIFS. Assume that the maps S_i are bi-Lipschitz for all i . Let $E \subseteq K$ with $\dim_P(E) < \dim_P(K)$. Then E is meagre in K .*

Proof This follows immediately from Lemmas 8.1 and 8.2. \square

Proposition 8.4 *Let X and Y be metric spaces. Let $f : X \rightarrow Y$ be a function and assume that:*

- (1) f is surjective;
- (2) f is continuous;
- (3) f is closed;
- (4) if U is a non-empty and open subset of X , then $f(X \setminus U) \neq Y$.

If X is not meagre, then Y is not meagre.

Proof Assume, in order to reach a contradiction, that the space Y is meagre. We can therefore choose countably many nowhere dense subsets N_1, N_2, \dots of Y such that $Y = \cup_n N_n$.

Since $Y = \cup_n N_n$ and $X = f^{-1}(Y)$, we immediately conclude that

$$X = \bigcup_n f^{-1}(N_n). \quad (8.5)$$

Next, we prove the following claim:

$$\text{the set } f^{-1}(N_n) \text{ is nowhere dense in } X \text{ for all } n. \quad (8.6)$$

In order to prove this, we let $x \in X$ and $r > 0$. We must now show that there is $x_0 \in X$ and $r_0 > 0$ such that

$$B(x_0, r_0) \subseteq B(x, r) \setminus f^{-1}(N_n).$$

Since the set $X \setminus B(x, r)$ is closed and the function $f : X \rightarrow Y$ is closed, we conclude that the image $f(X \setminus B(x, r))$ is closed and the complement $Y \setminus f(X \setminus B(x, r))$ is therefore open. Also, since $B(x, r)$ is open, it follows from assumption (4) on f that $f(X \setminus B(x, r)) \neq Y$. Consequently, $Y \setminus f(X \setminus B(x, r))$ is an open and non-empty subset of Y . In particular, it follows from this and the fact that N_n is nowhere dense that there is a point $y_0 \in Y$ and a positive real number $s_0 > 0$ such that

$$\begin{aligned} B(y_0, s_0) &\subseteq (Y \setminus f(X \setminus B(x, r))) \setminus N_n \\ &= Y \setminus (f(X \setminus B(x, r)) \cup N_n). \end{aligned} \quad (8.7)$$

By the surjectivity of f , we can find a point $x_0 \in X$ with $f(x_0) = y_0$, and since f is continuous we therefore conclude that there is a positive real number $r_0 > 0$ such that $B(x_0, r_0) \subseteq f^{-1}(B(y_0, s_0))$. We will now prove that

$$B(x_0, r_0) \subseteq B(x, r) \setminus f^{-1}(N_n). \quad (8.8)$$

Indeed, if $z \in B(x_0, r_0)$, then $z \in B(x_0, r_0) \subseteq f^{-1}(B(y_0, s_0))$, and it therefore follows from (8.7) that $f(z) \in B(y_0, s_0) \subseteq Y \setminus (f(X \setminus B(x, r)) \cup N_n)$, whence $f(z) \notin f(X \setminus B(x, r)) \cup N_n$, and so

$$f(z) \notin f(X \setminus B(x, r)), \quad (8.9)$$

$$f(z) \notin N_n. \quad (8.10)$$

We conclude from (8.9) that $z \notin X \setminus B(x, r)$, whence

$$z \in B(x, r), \quad (8.11)$$

and we conclude from (8.10) that

$$z \notin f^{-1}(N_n). \quad (8.12)$$

It now follows from (8.11) and (8.12) that $z \in B(x, r) \setminus f^{-1}(N_n)$. This completes the proof of (8.8) and hence the proof of (8.6).

Finally, it follows immediately from (8.5) and (8.6) that X is meagre contradicting the fact that X was assumed not to be meagre. \square

Lemma 8.5 *Let $(X, (S_i)_{i \in I})$ be an IIFS. Assume that $(X, (S_i)_{i \in I})$ satisfies the SOSC and that the map π is closed. If U is a non-empty and open subset of $I^{\mathbb{N}}$, then $\pi(I^{\mathbb{N}} \setminus U) \neq K$.*

Proof Since K satisfies the SOSC there is a non-empty, open and bounded set V with $K \cap V \neq \emptyset$ such that $S_i(V) \subseteq V$ for all i and $S_i(V) \cap S_j(V) = \emptyset$ for all i, j with $i \neq j$. Because $K \cap V \neq \emptyset$, there is $\mathbf{w} \in I^{\mathbb{N}}$ such that $\pi(\mathbf{w}) \in V$. In particular, this shows that $\cap_m S_{\mathbf{w}|m}(K) \subseteq \cap_m S_{\mathbf{w}|m}(X) = \{\pi(\mathbf{w})\} \subseteq V$, and since V is open and $\lim_m \text{diam}(S_{\mathbf{w}|m}(K)) = 0$, we therefore conclude that there is a positive integer m such that $S_{\mathbf{w}|m}(K) \subseteq V$. Putting $\mathbf{k} = \mathbf{w}|m \in I^*$, we therefore deduce that

$$S_{\mathbf{k}}(K) \subseteq V. \quad (8.13)$$

Next, since U is non-empty and open, we can choose $\mathbf{u} \in I^*$ such that $[\mathbf{u}] \subseteq U$. We now claim that

$$\pi([\mathbf{uk}]) \cap \pi(I^{\mathbb{N}} \setminus U) = \emptyset. \quad (8.14)$$

Indeed, since $[\mathbf{u}] \subseteq U$, we have $[\mathbf{u}] \cap (I^{\mathbb{N}} \setminus U) = \emptyset$, and it therefore follows from (8.13) and Lemma 7.2 applied to $H = I^{\mathbb{N}} \setminus U$ and $\mathbf{l} = \mathbf{u}$, that $\pi([\mathbf{uk}]) \cap \pi(I^{\mathbb{N}} \setminus U) = \emptyset$. This completes the proof of (8.14).

In particular, we conclude from (8.14) that $\pi(I^{\mathbb{N}} \setminus U) \neq K$. \square

Lemma 8.6 *Let $(X, (S_i)_{i \in I})$ be an IIFS. Assume that $(X, (S_i)_{i \in I})$ satisfies the SOSC and that the map π is closed. Then K is not meagre in K .*

Proof Note that the map $\pi : I^{\mathbb{N}} \rightarrow K$ is surjective, continuous and closed, and that it follows from Lemma 8.5 that if U is a non-empty and open subset of $I^{\mathbb{N}}$, then $\pi(I^{\mathbb{N}} \setminus U) \neq K$. Since, $I^{\mathbb{N}}$ is not meagre (for example, because the space $I^{\mathbb{N}}$ is complete), we therefore conclude from Proposition 8.4 that K is not meagre. \square

We can now prove Theorem 2.4.

Proof of Theorem 2.4 Assume, in order to reach a contradiction, that $\dim_P(\mathbb{E}_{\Pi}) < \dim_P(K)$. It follows from this and Lemma 8.3 that \mathbb{E}_{Π} is meagre in K , and since it also follows from Theorem 2.3 that $K \setminus \mathbb{E}_{\Pi}$ is meagre in K , we conclude that $K = \mathbb{E}_{\Pi} \cap (K \setminus \mathbb{E}_{\Pi})$ is meagre in K . However, this contradicts the fact that K is not meagre by Lemma 8.6. \square

Acknowledgements We thank an anonymous referee for carefully reading the manuscript.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix A

The purpose of this appendix is to prove (4.1).

Proof of (4.1) Fix an integer m with $m \geq 0$ and let $\mathbf{x} = (x_n)_n$ be a sequence real numbers with $x_n \geq 0$ for all n . Let $p_n = \frac{1}{n}$ and $q_n = \binom{n+m-1}{m-1}$ for $n \in \mathbb{N}$. Also, for positive integers n and i , let $P_n = \log(n+1)$, $Q_n = q_1 + \cdots + q_n$ and $u_{n,i} = \frac{p_i}{q_i} \chi_{[1,n]}(i)$, and define $a_{n,i}$ by $a_{n,i} = (u_{n,i} - u_{n,i+1}) \frac{Q_i}{P_n}$. For brevity write $C_n^m(\mathbf{x}) = \frac{1}{\binom{m+n}{m}} C_n^m(\mathbf{x})$. Since clearly $R_n(\mathbf{x}) = \frac{1}{P_n} (p_1 x_1 + \cdots + p_n x_n)$ and $C_n^m(\mathbf{x}) = \frac{1}{\binom{m+n}{m}} C_n^m(\mathbf{x}) = \frac{1}{Q_n} (q_1 x_1 + \cdots + q_n x_n)$, a straight forward (but lengthy) calculation shows that

$$R_n(\mathbf{x}) = \sum_i a_{n,i} C_i^m(\mathbf{x}). \quad (\text{A.1})$$

Finally, it is not difficult to see that the numbers $a_{n,i}$ are positive with $\lim_n a_{n,i} = 0$ for all i and $\lim_n \sum_i a_{n,i} = 1$, and it therefore follows from (A.1) and Toeplitz lemma that $\underline{C}^m(\mathbf{x}) = \liminf_n C_n^m(\mathbf{x}) \leq \liminf_n \sum_i a_{n,i} C_i^m(\mathbf{x}) = \liminf_n R_n(\mathbf{x}) = \underline{R}(\mathbf{x})$ and $\overline{R}(\mathbf{x}) = \limsup_n R_n(\mathbf{x}) = \limsup_n \sum_i a_{n,i} C_i^m(\mathbf{x}) \leq \limsup_n C_n^m(\mathbf{x}) = \overline{C}^m(\mathbf{x})$. \square

References

1. Albeverio, S., Pratsiovytyi, M., Torbin, G.: Topological and fractal properties of real numbers which are not normal. *Bull. Sci. Math.* **129**, 615–630 (2005)
2. Baek, I.-S., Olsen, L.: Baire category and extremely non-normal points of invariant sets of IFS's. *Discrete Contin. Dyn. Syst.* **27**, 935–943 (2010)
3. Barreira, L., Schmeling, J.: Sets of “non-typical” points have full topological entropy and full Hausdorff dimension. *Israel J. Math.* **116**, 29–70 (2000)
4. Billingsley, P.: *Probability and Measure*, 2nd edn. Wiley, Hoboken (1986)
5. Bugeaud, Y.: *Distribution Modulo one and Diophantine Approximation*. Cambridge Tracts in Mathematics, vol. 193. Cambridge University Press, Cambridge (2012)
6. Dajani, K., Kraaikamp, C.: *Ergodic Theory of Numbers*. Carus Mathematical Monographs, vol. 29. Mathematical Association of America, Washington, DC (2002)
7. Edgar, G.: *Integral, Probability, and Fractal Measures*. Springer, New York (1998)
8. Falconer, K.J.: *The Geometry of Fractal Sets*. Cambridge University Press, Cambridge (1986)
9. Hardy, G.: *Divergent Series*. Oxford (1949)

10. Hardy, G., Riesz, M.: The General Theory of Dirichlet's Series. Cambridge Tracts in Mathematics and Mathematical Physics, vol. 18. Cambridge University Press, Cambridge (1915)
11. Hutchinson, J.: Fractals and self-similarity. *Indiana Univ. Math. J.* **30**, 713–747 (1981)
12. Hyde, J., Laschos, V., Olsen, L., Petrykiewicz, I., Shaw, A.: Iterated Cesàro averages, frequencies of digits, and Baire category. *Acta Arith.* **144**, 287–293 (2010)
13. Jacob, M.: Über die Äquivalenz der Cesàroschen und der Hölderschen Mittel für Integrale bei gleicher reeller Ordnung $k > 0$. *Math. Z.* **26**, 672–682 (1927)
14. Kuipers, L., Niederreiter, H.: Uniform Distribution of Sequences. Wiley, New York (1974)
15. Lévy, P.: Théorie de l'Addition des Variables Aléatoires. Gauthier-Villars, Paris (1937)
16. Liao, L., Ma, J., Wang, B.: Dimension of some non-normal continued fraction sets. *Math. Proc. Cambridge Philos. Soc.* **145**, 215–225 (2008)
17. Madritsch, M.: Non-normal numbers with respect to Markov partitions. *Discrete Contin. Dyn. Syst.* **34**, 663–676 (2014)
18. Madritsch, M.: Non-normal numbers with respect to infinite Markov partitions. Preprint
19. Madritsch, M., Petrykiewicz, I.: Non-normal numbers in dynamical systems fulfilling the specification property. *Discrete Contin. Dyn. Syst.* **34**, 4751–4764 (2014)
20. Mauldin, R.D., Urbanski, M.: Dimensions and measures in infinite iterated function systems. *Proc. Lond. Math. Soc.* **73**, 105–154 (1996)
21. Mauldin, R.D., Urbanski, M.: Graph Directed Markov Systems. Cambridge Tracts in Mathematics, vol. 148. Cambridge University Press, Cambridge (2003)
22. Moran, P.A.P.: Additive functions of intervals and Hausdorff measure. *Proc. Camb. Philos. Soc.* **42**, 15–23 (1946)
23. Olsen, L.: Extremely non-normal continued fractions. *Acta Arith.* **108**, 191–202 (2003)
24. Olsen, L.: Applications of multifractal divergence points to sets of numbers defined by their N -adic expansion. *Math. Proc. Camb. Philos. Soc.* **136**, 139–165 (2004)
25. Olsen, L.: Extremely non-normal numbers. *Math. Proc. Camb. Philos. Soc.* **137**, 43–53 (2004)
26. Olsen, L., Winter, S.: Normal and non-normal points of self-similar sets and divergence points of self-similar measures. *J. Lond. Math. Soc.* **67**, 103–122 (2003)
27. Oxtoby, J.: Measure and Category. Graduate Texts in Mathematics, vol. 2, 2nd edn. Springer, New York (1980)
28. Riesz, M.: Une méthode de sommation équivalente la méthode des moyennes arithmétique. *C.R. Acad. Sci. Paris* **152**, 1651–1654 (1911)
29. Šalát, T.: Zur metrischen Theorie der Ljurothschen Entwicklungen der reellen Zahlen. *Czechoslovak Math. J.* **93**, 489–522 (1968)
30. Šalát, T.: Bemerkung zu einem Satz von P. Lévy in der metrischen Theorie der Kettenbrüche. *Math. Nachr.* **41**, 91–94 (1969)
31. Volkmann, B.: Gewinnmengen. *Arch. Math.* **10**, 235–240 (1959)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

L. Olsen¹ · M. West²

M. West
mawest@uci.edu

¹ Department of Mathematics, University of St. Andrews, St. Andrews, Fife KY16 9SS, Scotland, UK

² Department of Mathematics, University of California, Irvine, Irvine, CA 92697, USA